Supporting PhD physicists through investigations of ethical decision-making and structures of graduate education

by

Bill Bridges

B.S., University of Idaho, 2019

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

Doctor of Philosophy

Department of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2025

Abstract

The work represented in this dissertation is intended to support the development of PhD physicists in areas identified as critical to the PER community through established research in the community and from national reports, particularly in ethics education for faculty scientists and physics graduate education. I describe two individual projects that focus on considering different Epistemic Frames that faculty scientists engage in during value-laden discussions and the graduate education of physicists in terms of program requirements and retention measures.

Scientists engage with ethical dilemmas regularly. Training modules are developed to instruct scientists to better identify and navigate these dilemmas, but often focus on rules and regulations and not contexts aligned with the needs of scientists. Case-based discussions have been shown to impact ethical decision-making, and part of the Goals and Values in Science project was to investigate the impact that value-laden discussions have on faculty scientists' ethical decision-making. In this work, we applied the methodology of identifying epistemic frames to value-laden discussions. We identify three distinct group epistemic frames: Facilitator Frame, Sharing Science Frame, and Arguing Claims Frame. This work provides a foundation towards better understanding how faculty scientists make sense of value-laden discussions and promote productive avenues for designing ethical interventions.

The number of PhDs awarded in physics since 1861 has dramatically expanded, sometimes rapidly. To fully understand the potentially diverse set of program requirements and expectations of physics graduate education, we investigated the structures of PhD-granting institutions in the United States. We determined 10 distinct structures of program requirements, including Core Courses, Elective Courses, Comprehensive Content Exams, Research Readiness Exams, The Dissertation, Supervisory Committee, Training, Annual Reports, The Teaching Assistant Requirement, and Presentations of Research. These structures represent the most common practices and requirements of the physics graduate programs investigated. We additionally identified a common

narrative of physics graduate programs, the emphasis of physics content knowledge, and the diverse set of requirements for Candidacy Exams. The information from this study helps to support a foundation of current physics graduate program requirements and expectations, and how these structures could be further investigated to support students in physics graduate education.

The rate of physics graduate students awarded a PhD has historically been near 50%, with no consistent way to measure retention and completion rates for individual institutions. We used graduate enrollment survey data from the American Institute of Physics between 2000 and 2022. We propose 2 sets of measures of graduate student retention and completion rates that are consistent for different institution sizes and degrees awarded, allowing us to calculate retention and completion rates at both the national level as well as the individual institutions. The Graduate Retention Rates consider the flux of graduate student populations. We applied measures to the National Set of data that totals information from all 284 graduate institutions and then applied to the individual Institutional Level. We determined a high retention rate with opportunities to further refine the measure of retention. We identified an increase in the completion rates of all degree types at the National Level.

Supporting PhD physicists through investigations of ethical decision-making and structures of graduate education

by

Bill Bridges

B.S., University of Idaho, 2019

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

Doctor of Philosophy

Department of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2025

Approved by:

Major Professor
James T. Laverty

Copyright

Abstract

The work represented in this dissertation is intended to support the development of PhD physicists in areas identified as critical to the PER community through established research in the community and from national reports, particularly in ethics education for faculty scientists and physics graduate education. I describe two individual projects that focus on considering different Epistemic Frames that faculty scientists engage in during value-laden discussions and the graduate education of physicists in terms of program requirements and retention measures.

Scientists engage with ethical dilemmas regularly. Training modules are developed to instruct scientists to better identify and navigate these dilemmas, but often focus on rules and regulations and not contexts aligned with the needs of scientists. Case-based discussions have been shown to impact ethical decision-making, and part of the Goals and Values in Science project was to investigate the impact that value-laden discussions have on faculty scientists' ethical decision-making. In this work, we applied the methodology of identifying epistemic frames to value-laden discussions. We identify three distinct group epistemic frames: Facilitator Frame, Sharing Science Frame, and Arguing Claims Frame. This work provides a foundation towards better understanding how faculty scientists make sense of value-laden discussions and promote productive avenues for designing ethical interventions.

The number of PhDs awarded in physics since 1861 has dramatically expanded, sometimes rapidly. To fully understand the potentially diverse set of program requirements and expectations of physics graduate education, we investigated the structures of PhD-granting institutions in the United States. We determined 10 distinct structures of program requirements, including Core Courses, Elective Courses, Comprehensive Content Exams, Research Readiness Exams, The Dissertation, Supervisory Committee, Training, Annual Reports, The Teaching Assistant Requirement, and Presentations of Research. These structures represent the most common practices and requirements of the physics graduate programs investigated. We additionally identified a common

narrative of physics graduate programs, the emphasis of physics content knowledge, and the diverse set of requirements for Candidacy Exams. The information from this study helps to support a foundation of current physics graduate program requirements and expectations, and how these structures could be further investigated to support students in physics graduate education.

The rate of physics graduate students awarded a PhD has historically been near 50%, with no consistent way to measure retention and completion rates for individual institutions. We used graduate enrollment survey data from the American Institute of Physics between 2000 and 2022. We propose 2 sets of measures of graduate student retention and completion rates that are consistent for different institution sizes and degrees awarded, allowing us to calculate retention and completion rates at both the national level as well as the individual institutions. The Graduate Retention Rates consider the flux of graduate student populations. We applied measures to the National Set of data that totals information from all 284 graduate institutions and then applied to the individual Institutional Level. We determined a high retention rate with opportunities to further refine the measure of retention. We identified an increase in the completion rates of all degree types at the National Level.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I first need to thank my advisor J.T. for the years of support and mentorship. When I began grad school, I had never worked with physics education, having only learned it was a discipline the year before I started. The dedication towards helping me understand the work I would complete in these years is something I always appreciated. The weekly meetings were always a comfort and his greatest gift to me was creating an environment for me to ask questions and not feel lost in my work. Thank you J.T. for the years of stories, effort, and our collective ability to definitely stay on task.

Next, I would like to thank my committee members: Dr. Rachel Henderson, Dr. Scott Tanona, and Dr. Daniel Rolles. The help and feedback during my preliminary exam pushed me to better articulate my research understanding. Thank you Rachel and Scott for the years of feedback and support in working with you both on our respective projects. Thank you Daniel for being so supportive over the years, including my first time ever vising K-State.

Thank you Dr. Matthew Hedman for the thoughtful and productive introduction to physics research. It was a privilege to work on the data from Cassini. The chance to do physics research as an undergrad was a wonderful learning experience. Thank you Dr. Jon Herrington for the support in better understanding some elements of philosophy and psychology. At a time when the research was new to me in many respects, I always appreciated your support, knowledge, and good humor. Thank you Dr. Jackie Chini and Daniel Sharkey for the support over the years in making sense of the landscape study.

The beating heart of the physics department is Kim Coy. Thank you so much Kim for the tremendous amount of work you do for every student, the amount of care you show for all of us, and the many ways you have supported me over the years.

I am particularly grateful for the various cohorts of graduate students I have had the great joy in knowing and working with these last couple years. For Tyler and Shams, I appreciate our quick friendship, and the patience to answer my endless questions. It helped to provide a sturdy

foundation of support in 2020. Thank you Lauren for the many discussions ranging from books to crafting to musicals and your constant support in everything.

Maintaining good humor was critical to help mediate the stress of grad school for me, and luckily the last few years have had some wonderful people and experiences to help with this. At the end of this journey I want to also acknowledge those experiences and some quotes to illuminate just how silly grad school can be as well

- The Cincinnati Car Ride with J.T. and Tyler
- Some walks to the zoo
- A very ‘serious’ discussion of recycling at a conference
- ‘What, you don’t beep?’
- ‘The risk I took was calculated, but boy am I bad at math’
- ‘Graphic design is my passion’
- ‘Shorts don’t carry the weight of adulthood’
- ‘No! You CAN eat 20 hotdogs in 10 minutes, believe in yourself!’
- ‘I do not think, therefore I do not am’
- ‘Time is a soup, and I am a fork’

Dedication

To my Brother, Mom, and Dad, who is always with me

Chapter 1

Introduction

The development of a physicist involves many stages including undergraduate education and professional development. Instrumental to this evolving development focuses on improving this development and physics in general. The work represented in this dissertation involves two areas of physicist development that are active concerns in the community: ethics education of professional physicists and graduate physics education. A synthesis of Physics Education Research (PER) at the undergraduate level was published in 2014. Six topic areas of research, including problem solving, assessment, and attitudes and beliefs about teaching and learning, were covered in this comprehensive analysis. As described by the authors, this synthesis captures a good cross-section of PER, assessing where the research community has been and where it currently is. This synthesis is additionally non-exhaustive and does not capture the full scope of PER. This work also provides a perspective of what further investigations can be done and how PER can be expanded, broadening the range of student populations, and student reasoning and thinking [1]. In 2023, the American Institute of Physics published the International Handbook of Physics Education Research that includes chapters with these additional perspectives such as Research on Equity in Physics Graduate Education and the Aims and Values of Physics [2]. Other research involving the physics community is represented through national reports. The National Academies of Science, Engineering, and Medicine (NASEM) often publishes reviews of particular areas of STEM including the research integrity of students and faculty [3], graduate education and attrition [4;5], and a statistical analysis of doctoral recipients[6].

The national reports from NASEM help to highlight particular concerns in STEM, and while not an exhaustive list of the areas of PER, the synthesis from Docktor & Mestre[1] and the recent International Handbook of PER[2] both highlight particular research interests of PER, and the national reports from NASEM considers concerns of the STEM community. The work presented in this dissertation will consider research interests in PER that are aligned with the reports from NASEM and the International Handbook of PER in two particular areas. The first is through ethics education, and the impact that ethics training has on faculty scientists. A common training module for faculty and graduate students is ethics training, and present research considers the way that faculty scientists consider ethics as part of their research work is scarce. The second part of this dissertation concerns physics graduate education. So far, PER has engaged with graduate education at the scale of particular practices and elements such as admissions and candidacy exams, with organizations like the American Institute of Physics (AIP) and the American Association of Physics Teachers (AAPT) promoting further research into graduate education in general.

In this dissertation, I introduce two major projects that consider these two areas. First, I will describe my role in the Goals and Values in Science (GVS) project. GVS was designed to investigate how faculty incorporated values in their understanding of research practices. Motivated by the lack of effective ethics and Responsible Conduct of Research (RCR) training modules, this project investigated how value-laden discussions affected the fellowship scientists' ethical decision-making. The second project centers on two elements of physics graduate education. We investigated the diversity of program requirements of physics graduate programs across the country and identified the common structures. I additionally investigated the feasibility of developing measures of physics graduate student retention based on survey data from (AIP).

The organization of this dissertation is as follows: Chapter 2 summarizes the literature on education and training effectiveness, the construction of epistemic frames, and graduate education in the United States and the particular work of graduate retention. I describe my role in the GVS project in Chapter 3 and the group epistemic frames identified in the fellowship value-laden discussions. Chapter 4 discusses the methodology of identifying structures of graduate program requirements in the United States, and list the common requirements of each structure. In Chapter 5, I discuss the construction and application of yearly retention and completion rates of physics graduate students based on enrollment data from AIP. Chapter 6 discusses conclusions of both major projects while outlining places for furthering the work conducted in both.

1.1 The Goals and Values in Science

A survey completed in 2004 by Kirby & Houle revealed that 39% of the junior members of the American Physical Society (APS) reported observing or having personal knowledge of ethical misconduct, including reports of plagiarism, data falsification, and improper authorship[7]. A study of the prevalence of improper research practices also found that 15% of scientists have witnessed misconduct, while a further 40% have witnessed other questionable research practices[8]. Training in Responsible Conduct of Research (RCR) practices is essential for scientists in all disciplines, including physics. Recognizing the risk that ethical misconduct poses to the value of science, both the National Science Foundation (NSF) and the National Institutes of Health (NIH) have required all trainees to receive RCR training for more than 30 years[9]. The National Academies of Science, Engineering, and Medicine have released guidelines to enhance the relevance of RCR training[3;10]. Likewise, APS has also released guidelines to ensure that RCR training remains in line with the needs of scientists[11]. The dominant form of this training is self-paced, online modules (e.g., the CITI training portal)[12]. The survey from Kirby & Houle was administered again in 2020 to junior physicists and graduate students, showing mixed improvement. Most notable was that while physicists were being trained more often in ethics, the number of misconduct violations did not significantly change[13].

Further studies revealed the present method of training scientists to be largely ineffective and inconsistent[14]. One particular element of this training process concerns the ethical decision-making of scientists. Many tests exist to investigate and assess this, including the Defining Issues Test and the Ethical Decision-Making Measure. These tests present participants with ethical vignettes and ask them to rank the best choice of action for the subject from a short, defined list[3;15]. Studies of ethical decision-making have largely focused on the characteristics that impact it. One particular method considers how scientists engage with the key causes of ethical dilemmas in a case-based learning format, which results in better ethical decision-making[16].

This work is part of a larger Goals and Values in Science (GVS) project, interested in the ways that faculty scientists consider the values inherent in the research they conduct and the impact that discussions that center these values can have on their ethical decision-making. The judgments made in the progress of science require many considerations and decisions. We suppose that these decisions are made inherent to the values that the scientists have, where values represent the goals and principles that directly influence a decision someone makes[17]. This idea is at odds with the idea that science is inherently 'value-free', where scientists make decisions devoid of their biases[18;19]. As scientists make value-laden judgments at every level of the research process, denying the presence of these values denies the ethical responsibilities of those scientists[18;20]. The primary intervention part of the GVS project was a set of value-laden discussion sessions on a topic sitting at the intersection of science and society, such as the social epistemology of science. The sessions were constructed to allow participants to engage in discussions on the topic that would encourage the values of participants to be invoked.

Another study of the GVS project revealed that these value-laden discussions affect a scientist's ethical decision-making[21], and the part of the GVS project presented in this Dissertation is focused on how these discussions may have impacted the scientists of the fellowship. Focusing then on the value-laden discussions themselves, we propose a useful way to investigate how the scientists were making sense of the session is through epistemic frames. The methodology of identifying epistemic frames has been used to investigate the ways that someone may consider a discussion in one particular moment[22]. The framing experience itself explores the assumptions, skills, and tools that someone considers while engaging a particular task or situation[22–24]. In doing so, they necessarily limit the amount of relevant information, or resources, to just what they believe is needed for the task at hand. By observing the way someone behaves and argues, researchers have been able to identify the way someone makes sense of a situation. The collection of these pieces of resources, the behaviors, the warrants of the argument, and their resulting expectations form a single epistemic frame[24;25]. My role in this was to examine the ways the faculty scientists engaged in each discussion. We adapted the methodology of identifying Epistemic Frames to include the values invoked by the participants and the warrants used in each topic of discussion.

1.2 The structures of physics graduate education

Physics graduate education in the United States is historically complex and understudied. It is a system that has changed in process and scale since it began in 1861[26]. Graduate students have historically gained more specialized experience in research while training to become professionals in their field[27]. Concerns of the current system of graduate education include a weakening in the content taught to graduate students and that the current system is ill-suited to prepare students for careers after graduating[28–30]. This process is entirely based on historical context, with no overseeing organization to monitor student progress and establish a set of expectations and requirements for a Physics PhD. With each institution likely engaging in graduate education in isolation and potentially little interaction with the program requirements of other institutions, it is unclear how uniform the requirements of the degree are across the country. We are interested in the requirements of the graduate program and the potential variety of ways that a department keeps track of its students' progress. We consider these requirements of a physics PhD as part of the structure of graduate education. These structures highlight the distinct way each graduate program considers necessary to the physics PhD and what that program values.

Holistic reviews of graduate education are not common[31], and when conducted in the 1930s, the system of graduate education was originally found to be 'frustrating, confusing, and overwhelmingly difficult to describe at the national level'[32]. More recent surveys of graduate education in general have been conducted by the National Academies of Science, Engineering, and Medicine[4] and a joint task force of the American Association of Physics Teachers (AAPT) and American Physical Society (APS) have released report of graduate education in physics specifically[33]. These national surveys both considered a holistic review of graduate education while proposing recommendations. Significant work has also been done on particular elements of graduate education. These elements include determining what success means for graduate programs and students[34–38], the training of graduate students[27;28;34;35], and the retention of graduate students[39–45]. Within the Physics Education Research community, these investigations have surrounded three particular aspects of graduate programs: the admissions process[46], the perception of candidacy exams[47], and how students navigate through research groups[48].

We seek to expand research into physics graduate education to include the structures of program requirements across the country. With limited resources on how universal these requirements are, we seek to understand the potential variations of those structures. In understanding how uniform the program requirements of physics graduate education are, we can then invite a more critical analysis of the physics graduate system, including student retention, mental health, and other necessary aspects for furthering the experiences of students.

1.3 Measures of yearly graduate retention and degree completion rates

The retention of students is part of the concerns of graduate education listed above. Traditionally, these studies have focused on the retention of undergraduate students, resulting in a common understanding that 50% of students are being retained[45;49]. These retention studies have been instrumental in determining where students are having difficulties in their programs and finding ways to support undergraduate physics students in persisting. Considering these types of studies at the physics graduate level, little research has been done. Among the concerns of graduate student retention are both the impact it has on the students and the effect it has on the department. The graduate retention studies often propose a particular value, where the rate of PhD student attrition is between 40-50%[39–44;49].

The measures of retention often focus on cohorts of students at particular institutions or groups of institutions. In this way, the measure of retention/persistence/attrition of students can be informed by the reasons that students are making that decision[49–51]. More recently, the underlying reasons for students leaving have shifted to include the influences from the institution, and not only the students' academic experiences. These studies have been instrumental in understanding why students are leaving their programs and productive for identifying places of constructive changes in graduate programs[52].

We are interested in the potential connections between an institution's retention of graduate students and its program requirements. An exploratory step in this direction is to start investigating the changes in graduate student populations and retention rates at a national level for all physics graduate programs. In particular, we are interested in creating productive measures of retention and completion rates. The retention rate focuses on the yearly change in physics graduate student population, and the completion rate considers the rate of graduate degrees awarded to first-year graduate students. By introducing these measures, we aim to support institutions in reflecting on the retention and completion rates of their own graduate students to assess where changes can be made.

Chapter 2

Literature Review

This literature review will be split into two parts: first, to discuss the combination of research interests of ethics education, ethical decision-making, and Epistemic Frames in existing Physics Education Research. The second part will discuss the history of physics graduate education, including the long-standing concerns, present research, and finally, focusing on the retention of physics graduate students.

2.1 Goals and values in science

In the following section, we provide the landscape of current practices involving ethical training within the sciences and the scientific community's response to those practices. We then provide the background of student epistemologies and argumentation towards building epistemic frames in Physics Education Research.

2.1.1 Ethical training responding to misconduct

The concern of ethical misconduct in the scientific community is well known, particularly considering the impact that misconduct has on science, scientists, and the public[53]. Recent studies suggest that rates of questionable research practices and scientific misconduct are both high and considerably under-reported[54]. Where each action taken, choice, and judgment made by scientists carries the chance that they will lead to mistakes and false conclusions[55]. The National Academies also identified a need for improved training in research integrity[3]. The American Institute of Physics (AIP) updated its guidelines on ethics to further emphasize that part of a physicist's duty is to avoid ethical misconduct[11]. Kirby and Houle surveyed the junior members of APS originally in 2004, who responded with 39% having observed or had personal knowledge of at least one ethical violation in physics[7]. This survey was administered again in 2020 to junior physicists and graduate students, showing mixed improvement. Most notable was the idea that while physicists were being trained more often in ethics, the number of misconduct violations did not significantly change[13]. From the perspective of the scientists, Cairns et al. found that what scientists believe to be causes of unethical behavior match well with what other studies have shown to be actual causes of those behaviors[56].

The traditional state of ethics training in the scientific research community is largely focused on engaging with online training modules (e.g., The CITI Modules)[15;57]. These training modules are structured to present trainees with scenarios and vignettes depicting misconduct, including improper authorship or mishandling of data, and then being quizzed[56]. Other forms include a case-based ethics training that focuses on particular cases of misconduct, or through particular criteria that allow for a systematic and objective evaluation of ethical behaviors[58;59]. Federal agencies such as the National Institute of Health (NIH) and the National Science Foundation (NSF) have mandated the education of all trainees in the Responsible Conduct of Research (RCR) for years[54], with Federal requirements for RCR training established in the 1980's. The dominant mode of this training is through an online CITI Training System, which centers on trainees taking self-guided online modules followed by a quiz[56]. The NIH has additionally provided training grants for approved training programs[14]. AIP in 2019 put together a committee of ethics and produced a set of guidelines[11]. Additionally, the NIH has strongly recommended that training sessions include in-person discussions[14].

Concerns of effectiveness of ethics education

Evaluating the effectiveness of these various training programs is also well-established. Antes et al. conducted a meta-analysis of 26 ethics programs, showing that the overall effectiveness of ethics instruction was modest. More successful programs were case-based, interactive, and allowed participants to learn and practice the application of real-world ethical decision-making skills[60]; however, little is known about how specific case content influences training effectiveness[61]. Considering the current state of RCR training at each institution. There is a general sense that institutions have adopted the most convenient methods of satisfying RCR requirements, with little regard to their ethical impact[62], where there is little evidence that these standard models of RCR training are effective[12;53;56;60;63].

2.1.2 Ethical decision-making

Many tests exist that investigate and assess ethical decision-making, including elements such as the Defining Issues Test or the Ethical Decision-Making Measure[3;15]. These tests present participants with ethical vignettes and then ask them to rank the best choice of action for the subject based on their choice of action from a short, defined list. The response to these kinds of training modules is largely negative and often viewed as pointless[54]. The intention of these training modules is, in part, to impart an understanding of how to recognize and respond to various types of misconduct. These tests model a hierarchy of ethical behavior that subjects are more or less aligned with, ultimately assuming a model of deficit learning.

Studies of ethical decision-making have largely focused on the characteristics that impact it. One study evaluated personality characteristics that may have a positive or negative relationship with decision-making[64]. Another sought a causal analysis that could identify elements that forecast changes in this decision-making, highlighting that participants identifying critical causes of ethical dilemmas were associated with higher quality forecasting and 'higher ethicality of decisions'[65]. By outlining the key causes of dilemmas in a case-based learning format and emphasizing the negative outcomes, results in better ethical decision-making[16]. Antes et al. also found that by examining the effects of the existing RCR, the way participants respond to ethical problems involves an underlying process that impacts ethical decision-making[66].

2.1.3 Values in science

Hausman described a value in terms of a preference[17]. If someone prefers one thing over another, then that person values it. By preferring one thing to another, values must be considered as part of a relation. Values have particular relationships with one another, then, with some values conflicting (where someone may prefer two things but only one choice is possible) and others being supportive (where the end goal of one value inspires another separate value). Notable categories include ethical values, which include moral principles such as welfare or autonomy, or epistemic values, which involve the improvement of knowledge in some way. In a study by Linville et al., values were identified that were invoked while participants discussed research ethics. 3 distinct categories of values were identified that participants of this study would consider in the intersection of values within their research work[67]:

- **Ethical** - A value that appeals to what the participant perceives as the morally ”right thing” to do.
- **Epistemic** - A value that appeals to what the participant perceives as advancing or improving science.
- **Other** - A value that does not fall into the ethical or scientific categories. These include Legal, RCR, Economic, Communitarian, Self-Interest, and Practical.

The intersection of values in science has been studied for decades, and largely understudied despite its relative importance in the philosophy of science[18]. Investigating this intersection includes the direct role that values play in scientific practices. These practices include: framing the problem being investigated, the method of data collection, and understanding the results of a study[20]. The pursuit of understanding the impact that values have in science is something still being investigated[20;68].

Science is traditionally believed to be ‘value-free’, where scientists make decisions devoid of their own personal biases[18;19]. A defense of this idea can be found in the works of Du Bois, stating that public trust in science is dependent on science being without the values of those who pursue it[69]. While this notion is something to strive for, those conducting scientific research implicitly or explicitly make value-laden decisions. In another study, Elliot determined that decisions in science are inherently value-laden and require a number of different judgments. This study additionally determined that science cannot be value-free, with particular emphasis on particular arguments, including the conceptual argument (where non-epistemic values are necessary when assessing value-laden concepts) and the error argument (when scientists encounter difficulties in studies, they should consider non-epistemic values when making decisions)[68].

2.1.4 Resource framework

The concept of resources sits at the intersection of cognitive science, education research, and behavioral science that evolved from diSessa's framework of Knowledge-In-Pieces[22;70;71]. A resource is a small, reusable piece of knowledge and information that acts as a cognitive modeling element[72]. These epistemic resources represent a tightly bundled packet of information that, when activated, leads the individual to interpret the knowledge at hand in a particular way. Resources are dynamic, activated, and deactivated moment-to-moment during an activity[22]. For example, a student working on a physics problem for the first time that discusses forces would implicitly activate resources related to the student's present understanding of forces. The activation, suppression, and formation of the network of resources is dynamic as the action related to the resource is dependent on the present context[22;70]. For physics students, the activation of these resources is 'local', where the particular resource is activated within a network of many epistemic resources to productively solve problems and make sense of a phenomenon[22;25]. Representations of these networks of resources have been investigated in terms of the effect of particular resources being activated and the effect it has on the structure[73].

2.1.5 Identified warrants in arguments

The field of argumentation and rhetoric focuses most on presenting an argument[22;74]. The significant work of Toulmin identified a system of parsing elements of argument into distinct parts: claims, data, and warrants[75]. A person will make a statement, the *claim*, which requires proof.

Relevant facts, or *data*, will then be presented to support the claim. The *warrant* then is the element that bridges the supporting data with the claim[22]. Argumentation plays a central role in the building of explanations, models, and theories as scientists use arguments to relate the evidence they have to the claims they reach through the use of warrants and backings[76].

Bing & Reddish explored the presence of warrants in the ways that students worked and argued through a problem. Several notable elements of this work that involve warrants include how students' warrants can shift from moment to moment and how these shifts are closely tied to what they interpret as the nature of the current activity. The different proofs of arguments used during these shifts provide different ways that students may be framing each discussion of the same activity. The warrant itself then shows what epistemic assumptions are being made in that particular moment of the argument or activity. Therefore, the use of the term warrant focuses the attention on the epistemic nature of the argument being made, re-framing the warrant as an observable epistemic resource[22].

2.1.6 Epistemological framing

The dynamic and extended process of how someone makes sense of a particular situation or action is referred to as framing[23;77–80]. The term 'framing' is based on previous works through sociology, sociolinguistics, and cognitive science[78;79]. This process is colloquially referred to as' what is it that's going on here?'[22;24]. The challenge of understanding an individual's thinking includes a diverse range of conceptual and epistemological possibilities available. Modeling this process becomes manageable if the model can constrain this range to a manageable set of considerations possible to the individual during that time[22–24]. This sociolinguistic and anthropologically specific variation of framing accounts for a more dynamic process on the time scale of minutes to hours[81]. The focusing of particular cognitive elements additionally considers the skills and tools necessary for the present situation[22;23;25;77]. As described in Section 2.1.4, resources are small cognitive elements, formed in networks, activated by the particular context[22;73;82]. This process of selecting particular information and skills, or cueing, is informed by an adaptive network of cognitive resources, where shifts in framing are then accompanied by cueing different epistemological resources[22]. Frank & Scherr investigated a group of students' collective framing experience. This study found that the collective of individual students' micro-processes helps to sustain student thinking during a collaborative learning activity. They also found evidence that each student makes sense of the situation at various times, and often does so together. Particular features of the context contribute to the activation of the knowledge for each student[80].

Investigating epistemic framing directly is difficult. Instead, particular characteristics are used to infer what the framing experience is for individuals. Previous studies identified characteristics such as body language and meta-verbal cues (e.g., tone of voice, volume of speech, and rhythm in turn taking)[25;78;80]. For students working in groups, the characteristics identified shed more light on the nature of framing itself and the kinds of signals that groups use to communicate with one another during a group activity[25;78;83]. The method used by Scherr & Hammer determined discrete categories of characteristics used to investigate epistemic framing. This then implies that there are distinct boundaries between epistemic framing experiences where the framing experience can be external to one individual[24;25]. More work is necessary to determine what can cause a shift in framing, and the manner of the change itself[25].

2.1.7 Epistemological frames

The epistemic framing experience incorporates multiple elements, as outlined above in Section 2.1.6. The Epistemic Frame is then the complete set of expectations and activated resources based on an individual's framing experience[24;25]. These frames can be envisioned as a storage area for conceptual and procedural resources, which promotes the activation of particular epistemological resources and blocking others[77]. As someone makes sense of 'what it is that's going on', they implicitly limit the number of necessary resources based on the expectations they have for the present situation. Epistemic Frames exist at both the individual and group levels, which coincide with individual and group epistemic framing experiences. Bing & Reddish found that students can get trapped in Epistemic Frames for 'many minutes', finding them to be 'sticky'[22]. This attribute of Epistemic Frames can be potentially problematic for students existing in inappropriate Epistemic Frames for particular problems[70]. The characteristic markers used to identify these frames similarly come from linguistics[78]. In the particular space of education research, markers used to identify frames have historically been focused on student behaviors and verbal cues[24;25] and the substance of student arguments[25;84]. Irving et al. observed students' actions (and co-occurring clusters of actions) in order to infer what their frame might be[25].

2.2 Graduate education in the United States

The first PhD was awarded in 1861. There was a push in post-secondary education to try and expand doctoral graduate education beyond the disciplines of Theology, Law, and Medicine. In the formation of the Doctor of Philosophy degree, graduate students in 1860 at Yale University engaged in several topics 'beyond the undergraduate level'. Physics was one of the three PhDs awarded the following year. Several attempts were made in the preceding years to institute a doctoral program for those disciplines. The primary vehicle for these programs started in the advanced courses that would be offered, with some institutions such as Harvard, Brown, and Princeton proposing graduate courses as early as 1829. The graduate program established at Yale additionally required the writing and defending of the doctoral dissertation. The dissertation itself was described as an 'exercise in scholarly research', and symbolizing 'the research ability of the doctoral candidate in independent and original research'. Beyond these light descriptions, not much explanation has been given on the development and justification of the inclusion of the doctoral dissertation[26]. This new system of graduate education was heavily modeled on medieval universities and the German research institution model of the early 1800s[85].

Since 1861, the number of PhDs awarded has been steadily and aggressively increasing. In the 1920s, the number of PhDs awarded nearly tripled, increasing by another 50% in the 1930s. The Association of American Universities began working in the 1900s towards standardizing the requirements of a PhD, ultimately finding the increasingly varying doctoral programs difficult, and few efforts were made[32]. A Carnegie Foundation researcher in the 1930s found graduate education 'frustrating, confusing, and overwhelmingly difficult to describe at the national level', concluding that individual departments held sole responsibility for standards[32]. In the early 20th century, several significant changes had taken place in the space of graduate education. The university structure was forming as the only venue for professional credentialing for many fields[29]. This credentialing and research have driven the development of the system of higher education without requiring any central organization of that system[29]. The qualifying exams were implemented to address concerns that the defense of the dissertation, 'had decayed into an exercise serving only a ritualistic function'[86]. To further highlight the comments of the Carnegie researcher, graduate education is characterized as decentralized, competitive, and stratified[29], with no governing body that oversees the conduct of graduate education above the level of individual departments and university administrations[29]. Also in the 1930s, practices including the GRE and the use of comprehensive candidacy exams also became more common practices in graduate programs. In the mid-1990s, the number of graduate students in the science and engineering fields had declined. This number has since been consistently rising in the 21st century, and then reaching an all-time high[42]. During this time, a trend of degrees awarded to graduate students also changed, where the number of Master's degrees awarded between 1995 and 2005 increased by 47%. The number of doctorates awarded increased by 26%[42].

Concerns of the system of graduate education

Concerns of graduate education are nearly as old as the system itself. In an essay originally published in 1903, William James critiqued the reliance on the PhD as a credential necessary for teaching in the 'PhD Octopus' calling the practice 'a sham, a bauble, a dodge, whereby to decorate the catalogues of schools and colleges'[85]. With no central organization of graduate education, and rapidly expanding in the 19th and 20th centuries, the system of higher education was quickly characterized as decentralized, competitive, and stratified[29]. Specific concerns of graduate education involve several elements of the system. In particular, these concerns largely concern the rapidly changing job market and the overall quality of education. The 'job crises' of PhD students is well documented, with multiple facets. First is that there are more graduates than academic and research jobs[28–30]. The specialized training that graduate students receive prepares students for careers in academic and research roles, and is often ill-suited towards graduates taking jobs outside these roles[28;30]. Additionally, students finding positions in these roles are finding the transition between graduate education and careers to be difficult[28]. Considering then the quality of graduate education itself, where prominent concerns in the 19th century over weakening standards in awarding graduate degrees, the poor perceptions abroad of US research universities, and a steady migration of US graduate students to Europeon institutions led to the founding of the American Association of Universities in 1900, which found the process of finding a standardized organization of graduate education too difficult[85].

Reviews of the system

Holistic review of graduate education is not a common practice[31], yet it has been proposed that graduate education exhibits certain system features on a national scale. Credentialing and research have driven the development of the system of higher education without seeming to require any central organization of that system[29]. With no governing body overseeing the conduct of graduate education, faculty and directors often know little about what is happening at other universities[29;87]. In 1960, Berelson's Graduate Education in the United States offered a systemic review of graduate education that posed fundamental questions of what a PhD means and how quality and standards can be maintained[85]. Investigating graduate education itself as a whole has been shown to suffer from a lack of organization and definition, as the system is too complex and variable[35;44]. Reviews and research into graduate education have considered the system as a whole and individual parts, informing the whole. Historically, many studies have focused on graduate education as a whole as opposed to these individual elements[88]. Holistic reviews of graduate education include assessing outcomes and creating stronger connections between experiences in graduate education and the employment future of students[85], an evaluation of student learning and performance[35;89], considering any predictive nature of admission metrics[34], and identifying the goals of graduate education[35]. In the physics graduate research community specifically, efforts have been made to discuss trends and practices in physics graduate education[90]. Identified changes in physics graduate education based on these reviews and efforts include the AAPT releasing a statement advocating against the use of the Physics GRE in admissions, and APS launching the APS Bridge Program and the IGEN[91].

The National Academies: Graduate STEM Education for the 21st Century

One particular holistic review of graduate education includes The National Academies of Science, Medicine, and Engineering analysis of graduate education in STEM done in 2015[4]. The analysis was ordered in part to respond to the concern over graduate programs not adequately preparing students for careers beyond graduate school and, 'how well the current STEM graduate education system is meeting the full array of 21st century needs'. The report from this analysis calls for a systems approach to moving graduate STEM Education forward, and for each element to help achieve an ideal version of this system. The report particularly highlights that graduate education should shift to a more student-centered approach, where greater emphasis and focus is placed on graduate students as individuals with diverse needs and challenges. A critical element is the report's articulation of the core competencies that all students should acquire.

2.2.1 Defining graduate education and success of programs

Determining the success of graduate programs is ill-defined[35]. Several attempts have been made to determine the outcomes of graduate education, including elements such as objective and subject criteria including grades, publications, and predictions of future success[34]. An increased attention to these outcomes has been promoted by academic researchers, education-related foundations, and graduate students[85]. Researchers have tried to determine a more comprehensive definition of success in graduate programs and what elements contribute to that definition[34;35]. Attributes contributing to this type of definition fall into two groups: institutional contributions and student influences.

The success of its graduate programs is naturally varied from program to program depending on the discipline and institution[36]. The goals can vary as programs train researchers, teachers, and practitioners. The way each program determines success must be based on what skills are necessary for each role the graduate students take after leaving grad school[35]. A traditional sense of the health of a graduate program has been determined by faculty productivity, and interest in student-specific indicators of success in graduate programs has recently become more prevalent[85].

There are numerous attributes of students that are related to their graduate education and the way they contribute to the outcome of graduate education[36;37;89]. For many researchers, the training students receive in research plays a central role in the larger definition of graduate education[28]. The traditional view of graduate education is very research-focused, where students become researchers who produce new knowledge[35]. Faculty identify other attributes as high grades, a reasonable time to degree, and a student's ability to find a job post-graduation[85;92–94]. Another set of attributes includes the careers students are in years after graduating, and whether their graduate experience properly prepared them for that career[85]. Other notable elements of graduate education also include admission, the GRE, retention, and GPA[92;93].

2.2.2 Focus on particular practices and requirements

The complex system of graduate education contains numerous practices, including the dissertation, courses, and candidacy exams. As stated, there is no organization to standardize the experience of graduate education. As the subject matter of different departments and disciplines is very different, the faculty in each department naturally want to determine the requirements of a degree to reflect the training they received in that particular discipline. Many institutions have also recognized the need to have some consistency across departments, allowing for graduate colleges to establish particular requirements such as minimum admission standards, credit hour requirements, and residency requirements[95]. Another consistent element of practices in graduate education concerns certain 'benchmark assessments', including the qualifying exam and the dissertation defense that all graduate students experience[31].

Concerns of practices

Concerns of graduate education are also directed at particular practices, including the critique of the PhD from 1903[85]. The known milestones are often the subject of these concerns, with a particular emphasis on admission metrics, candidacy exams, and the dissertation. Admissions to graduate programs have predominantly been based on measures of classroom performance or tests that predict classroom performance, with groups suggesting that metrics should investigate evidence of potential 'professional distinction' in situations outside of testing and classroom performance[35].

Significant research has been conducted on the impact of the process of candidacy, involving both a comprehensive exam and a qualifying exam. in 1932, Pressey et al. published *The Final Ordeal*, which considers the role that this practice has on both students and the programs implementing these types of candidacy exams. Both the oral examination and final comprehensive written examination are thought to be trustworthy bases for decisions that vitally affect a candidate's education and professional future. The written comprehensive examination is not well-balanced or well-proportioned, and appraised in a rough and untrustworthy fashion, while the oral examination serves a definite purpose as an occasion for group discussion of candidacy, it is not viewed in this way. Evidence of the oral examination finds it to be distinctly unreliable, and does not show the validity as hoped by its implementation. The view of these exams is largely negative, as the exams are seen as serving a policing function. The oral examination, in particular, is seen as little more than a trial by combat, and an excuse for failing a poor student when the major professor cannot find it possible to do so. The authors of this work recommend that departments that careful consideration in administering these exams to holistically appraise the work of a graduate student[96]. These sentiments continue in more recent research endeavors. Liera et al. in 2023 found the doctoral qualifying exams and other candidacy requirements as gate-keeping practices that impact degree completion, equity, and student well-being[86]. A growing number of faculty, staff, and students in the physical sciences recognize that traditional qualifying exams deviate from the work expected of scholars and criticize them for low passing rates, unrealistic time constraints, unclear expectations and procedures, and an apparent apathy among faculty to mentor and prepare students[47;86]. The belief that comprehensive and qualifying exams are a policing measure used to screen out students perceived as weak by the program[97]. The process of candidacy is particularly difficult for underrepresented groups in physics and also viewed as a gate-keeping practice[98].

2.2.3 Training of graduate students

The training of graduate students in research is a central pillar of doctoral education[28]. Graduate school gives students the opportunity to select a field of study, a topic of interest, and a chance to investigate a relevant question in the field[27]. As students take careers in various roles after graduate school, successful training of those students towards these roles is determined by very different skills[35]. The way that these skills are taught manifests in different ways across graduate education, and generally, the way graduate programs are preparing students for many of these careers is in question[34]. There is a strong emphasis on students acquiring particular skills in writing, argumentation, and other evaluations of research. Many of these skills are not formally taught to students, and instead potentially acquired indirectly through research experiences. Students would additionally be modeling and demonstrating these skills during lab meetings and seminars. If we consider the training of graduate students as predictors for professional success, several elements of graduate education, such as identifying a problem for investigation and critiquing the work of collogues, can appear as part of the Master's Thesis or Doctoral Dissertation. As such, there is a lack of standardized models of assessing these skills and characteristics, likely owed to the fact that there is a lack of consensus on what constitutes the explicit goals of graduate education[35].

Another avenue of investigation of the training of graduate students is found after they have already left their programs. Traditionally, the outcomes of graduate education are bound by the years the student is actively engaged in graduate education. These longitudinal studies have been critical in pushing past degree completion in understanding the career choices of the graduate students and how their graduate education impacted those career decisions[85]. With an established 'job crisis' of few academic careers for the multitudes of graduate students, the overall recommendation from several studies is that graduate education should adapt to alternative career paths and provide versatile training for these paths and information on different options available to students[28;42]. These versatile training opportunities could include helping students develop explicit skills in project management, understanding of finance, funding and resource management, professional and research ethics, and relating the application of research to a more social context[42].

The socialization process

The traditional pedagogy of training graduate students is based on a discipline-centric relationship between student and advisor within a focused community[31]. In research-oriented programs, this training model can be viewed as a process of academic socialization, where students are expected to move from apprenticeship to independence. More successful modeling of this process usually occurs in environments where the faculty are actively engaged in research activities[35]. This socialization process is a transformative process where a novice identity is converted into a professional one that encompasses newly acquired specialized knowledge, skills, techniques, values, and attitudes[27;99;100]. Experts are believed to have a great deal of domain-specific knowledge that is hierarchically organized. Success in graduate education, then, is partially defined by perceived professional success, where students exhibiting behaviors and skills reminiscent of a professional are highly favorable. These professional skills include writing, argumentation, and evaluation of research, many of which are not formally taught by graduate programs and are acquired by students informally. Part of this includes practical professional knowledge, and for many students, the only opportunity to demonstrate and model these skills is in lab meetings and seminars. This modeling may also occur in graduate courses where students learn the ways in which leaders in particular fields may consider problems and issues[35]. In the physics community, the model of a physics expert is traditionally defined by the cognitive abilities of experts, such as having large amounts of content knowledge and having superior memory retrieval mechanisms for problem solving developed through ten years of deliberate practice. For physics students, this process may begin as early as the first introductory physics course in high school or college, but likely happens during graduate school, where students would have opportunities to develop skills and knowledge in a supportive community through deliberate practice and physics research[27].

Experiences and concerns from graduate students

A critically under-researched element of graduate education is the experiences and perceptions of the graduate students themselves[43;85]. Understanding this experience and perception involves multiple parts, including the support that students receive as graduate students as part of their changing identity[27]. Understanding these lived experiences of graduate students is important in understanding the barriers that may exist and identifying distinct ways of supporting students[41;43;99;101;102].

The student response and perception of the training they receive highlight some potentially problematic practices in graduate education. Several areas were specifically mentioned, including the preparation of academic work, preparation for careers outside of academia, the development of scholarly competencies, involving conceptual understandings of a particular field, knowledge and skills relevant to their area of study, interpersonal skills, and professional attitudes and beliefs[99]. In a survey of 4000 doctoral students in 2001, students reported that the training they received in their programs was not what they wanted and did not prepare them for the jobs they would eventually take[30].

2.2.4 Physics graduate education research

The numerous studies listed above concern many different aspects of graduate education dating back several decades. Many of these studies considered graduate education from an interdisciplinary perspective, and no study considered physics solely for research. Within the space of Physics Education Research, the interest in graduate education has been largely focused on particular aspects, and little about the system of physics graduate education as a whole. For instance, Young et al. found that allowing for a more holistic review process of admissions to graduate programs allowed for a more diverse and equitable graduate student population[46]. More recently, there has been a shift in focus beyond admissions studies, moving more into physics student perceptions of candidacy exams[47] and students having difficulties in research groups[48]. With these studies, we see this continued emphasis towards understanding the precise difficulties in the structure of graduate education. With an emphasis on better defining the problems in physics graduate education, this is not only a concern of various research groups but also organizations in physics education.

A task force to study the status of graduate education in physics in the United States was formed between the American Association of Physics Teachers and the American Physical Society in January of 2003[33]. The standing Committee on Graduate Education called for this task force partly in response to discussions of an expressed concern that changes in graduate education would weaken the ‘unity’ of physics as a discipline, echoing the concern listed above. These changes were listed as fewer common courses required and lowering the level of the comprehensive exam. Another reason for this report is that physics as a discipline has hardly changed in the previous 50 years.

The task force itself would then be charged with identifying the following points: the current status of graduate education, the graduate student experience, and departmental issues. The authors proposed at the time that the central curriculum should cover some form of Classical Electrodynamics, Quantum Mechanics, Classical Mechanics, Statistical Mechanics and Thermodynamics. They additionally recommend various training programs, such as communication and literacy, and that programs take a more active role in monitoring student progress to the PhD.

2.2.5 A system deficit model

The traditional effort to understand student difficulties in STEM courses is framed as, 'understanding how [underrepresented students] managed to acquire various forms of capital they did not possess upon entry' [103;104]. This can be described as a student-centered deficit model [105;106]. Harper proposes that instead of focusing on models that persist deficits, researchers examine an anti-deficit model to determine, 'what do students of color have', rather than 'what do students of color need?' [103;104]. This model is informed by psychology, sociology, and education and is built to shift the focus from the needs of the students to what students are using to support themselves. The model then shifts from the student-centered deficit model to a structure-centered deficit model, where asset-based approaches can promote learning in meaningful ways [103].

The application of this model has been investigated on several scales. Cotner & Ballen determined that several features of a classroom environment can form barriers and create a deficit in the performance of underserved students. These barriers were also shown to be mitigated by instructional choices and removal of particular elements such as 'cues that foster stereotypes' and 'use of role models as examples' [107]. Webb & Paul determined that centering a course-deficit model for an introductory physics course appropriately closes some demographic grade gaps [105]. Later work by these same authors highlights that differences in performance between different student groups in a given course to the features of the course itself. The course deficit model re-frames equity gaps as a product of the course that is failing to leverage the assets of a diverse student population, rather than blaming the equity gaps on the students. With this approach, 'the system must adapt to the changing population because the alternative is to force our students to adapt to an outdated, inequitable system'[108].

2.3 Retention of graduate students

Concerns of student attrition and retention are well-established for undergraduate students. Models have been established for years that build on the work of Tinto[109], where developing a model of student retention involves considering other elements. This can be seen in the work of Bean, who broadly sought to investigate the determinants of student attrition[49], and Berger & Milem, who worked to further the understanding between behavioral involvement and perceptual integration in college persistence[50]. This highlights a trend in developing models of attrition and retention, where additional elements inform those models. To accomplish this, conceptual models were made at single institutions, employing follow-up surveys to attempt to establish why students leave an institution[49]. These conceptual models have been applied to undergraduate students at various institutions involving both traditional and non-traditional students[110]. These models involve a path analysis technique that examines the direct and indirect effects of variables in these longitudinal studies of persistence[50]. A common problem with these previous models of student attrition lies in the fact that the definition of variables used in the analysis rendered the models unsuitable for path analysis[49].

The seminal works of Lovitts built upon the work of Tinto[40;109] to consider retention at the graduate level. In doing so, similar methods of conceptual modeling and path analysis have been considered at the graduate level. Doctoral attrition has been considered important for 3 particular reasons: first, that attrition at the doctoral level is poorly understood; second, consistently high levels of attrition may signal underlying problems in a department, university, or discipline; and third, that research on attrition is motivated by the economic and psychosocial cost of attrition. The modeling of undergraduate retention has informed much of early models of graduate retention, with little theorizing of models distinct from these modeling techniques[39]. Doing so has allowed researchers to determine the importance of graduate institutional characteristics and the impact they have on positive student outcomes, where these characteristics may act as barriers for students[39;51]. 30% of the reasons that students noted as contributing to their attrition belonged to departmental reasons such as inadequate advising, financial support, and departmental policies [41;111;112]. Studies like these helped expand reasons for students not persisting to include departmental and institutional factors rather than solely the students' academic experiences [111]. Lovitts contributed to this idea by pointing out that 'the reasons for attrition are embedded in the organizational culture of graduate school and the structure and process of graduate education' [40].

Models of attrition, retention, and persistence often follow the establishing techniques of path analysis and longitudinal models [51;113]. These models were determined on several scales, constructed based on single institutions to nationally [42;51;94]. A product of these conceptual models considers values of retention, attrition, and completion or graduation rates. Lovitts concluded that the graduation rate for PhD candidates is as low as 50% [40]. Other studies have found similar results. Golde found that 40% of the students who began a doctoral program fail to complete it [39], Terrell et al., who determined the attrition rate of doctoral students to be 40 to 50% [41], and the Ph.D. Completion Project determined cumulative completion rates to be 58% for men and 55% for women [94]. While 50-60% remains the more common rate of completion [39–41;94;114], Wendler et al. also determined non-completion rates at one institution to vary dramatically by discipline, ranging from 29-63% [42].

2.3.1 Retention studies in STEM

Now, considering retention and completion rates of STEM students, these modeling techniques persist. A push to understanding retention has taken more significance when considering that enrollment in higher education may be facing an enrollment cliff [108]. In continuing the trend of conceptual modeling, Seymour & Hunter determined a hierarchy of 23 factors that contributed to students' decisions not to persist in their STEM majors [115]. This technique has been used for various student populations and the effect different characteristics have on those students persisting in their STEM programs [115;116]. In physics, this modeling has considered the physics students and curricula of the department [117].

Similarly, transitioning to graduate programs, conceptual models have been used to understand factors that contribute to attrition in graduate programs. Owens et al. additionally found that the visions of faculty and graduate students were not aligned despite the two groups being aligned in improving retention[52]. Particular characteristics identified as having a significant impact on student graduate completion include the impact of the GRE on student completion rates, resulting in higher GRE scores being more likely to depart[111]. Graduate retention and completion rates have been determined to be harder to study, citing a number of issues, including the lack of uniformity in student populations, program requirements, definitions of PhD students, and undefined time to graduation. Gross et al. additionally determined that the completion rate for PhD students is at or slightly above 50%. The balance is not made up by students who receive a Master's degree. This suggests that roughly 25% of the students who enter graduate school leave in some 'undocumented way'[114]. Maher et al. determined the overall attrition of graduate students in a medical school to be 5.7% through a cohort analysis lasting 10 years[118]. Owens et al. found retention rates for under-served groups in graduate physical science and mathematics programs to be around 50%[52], while Sachmpazidi & Henderson determined retention rates for graduate students in science and engineering doctoral-granting departments to be 59%. Concerns of attrition in STEM graduate programs share those listed above, with particular emphasis on the effect on underrepresented groups and the emotional and financial cost of attrition[111].

2.3.2 Physics graduate retention

The modeling of graduate completion rates has similarly been conducted in physics graduate education. This has illuminated factors that influence doctoral completion and attrition, including disciplinary issues and mismatched expectations between students and their departments, poor advising, and structural isolation of students. These modeling techniques additionally highlight that while research on attrition is prevalent, and resulting improvement on placement tests, communication skills, mentoring of students, professional development opportunities, and productive participation in the department; attrition in graduate school remains high, with nearly half of the students enrolled in graduate PhD programs do not graduate with their physics PhD within 9 years[27]. Further investigations are needed to fully understand the factors that impact attrition, and it remains an important priority for researchers. Graduate attrition in physics continues to be a multifaceted phenomenon, largely not well-understood, including a lack of a systematic way of recording the reasons that lead students to depart their programs. The Bridge Program has developed several key recommendations to address this, including mentoring and research engagement, and better support for students. Students affiliated with the Bridge Programs have been shown to have a 92% retention rate [111].

2.4 Summary of Literature Review

Ethics training and decision-making

Ethics training for scientists has been required by federal agencies for decades [56]. Despite this training, the number of ethical misconduct reports is still significant [7;13]. This has called traditional ethics training, including RCR, into question. With more recent emphasis on someone's ethical decision-making, case-based training formats are productive for impacting someone's ethical decision-making. In the fields of anthropology and sociology, an epistemic frame is a way to capture someone's particular way of making sense of something. The methodology of identifying epistemic framing that results in an epistemic frame can be characterized in several ways in Physics Education Research, notably through student behaviors and warrants through a framework of epistemic resources.

Physics graduate education

The first physics PhD was awarded in 1861, and the system of graduate education rapidly grew in the following years. As the system grew, practices were added, such as candidacy exams, as deemed necessary because of concerns in the then-present system. A central organization of graduate education does not exist, and a product of this was a system already considered to be complex and difficult to describe at a national level by the early 1930s. Research into graduate level is a newer area of physics education research, and several critical elements that promote changes in graduate programs include admissions and the process of candidacy. Holistic reviews of graduate programs are not common, but many elements of graduate education have been reviewed.

This includes understanding the success of graduate programs, the impact of particular practices on student success, and the training of graduate students as researchers. Structural analysis of graduate education illuminates a path towards understanding the direct impact that different structures have on student success.

A particular concern in graduate education is student retention. Understanding the barriers and elements of the university system that affect students persisting in the program has helped reform elements to improve student retention. Student retention is well-documented for undergraduate students and is often found by modeling particular cohorts of students. Similar studies have been conducted for graduate students, often focusing on particular cohorts of students at single institutions or multiple institutions. At the graduate level, the common conception is that nearly 50-60% of graduate students enrolled will be awarded a PhD. The models used to determine the results are often through keeping track of cohorts of students or individual students' particular pathways and the particular parts of graduate education that are pushing students out or keeping them in.

Chapter 3

Identifying epistemic frames in value-laden discussions

3.1 Motivation and outline of this chapter

The study presented in this chapter represents a growing interest in the ethical responsibilities of scientists. The history of training in ethics and RCR is mired by ineffective modules and practices, and the number of ethical violations has persisted despite decades of training opportunities. A more recent approach is aimed at understanding the ethical decision-making of scientists and how this affects the way that they navigate the difficult decisions made in conducting research. The methodology of identifying epistemic frames has been used to make sense of how students engage in problem-solving as a group, and we propose this as a promising start to understanding how faculty scientists navigate discussions that center their own values. In particular, this chapter will be guided by the following research questions:

- What group epistemic frames are present in value-laden discussions?
- What group epistemic frames are characterized by the values invoked by faculty scientists?

Section 3.2 will establish the contributions of ethical decision-making and values in science towards defining the epistemic framing experience and frames of these value-laden discussions. The adapted methodology of identifying these frames will be described in Section 3.3. The 3 distinct group epistemic frames outlined in Section 3.4 will be contextualized in terms of established group frames, and the value they bring to understanding ethical decision-making will be discussed in Section 3.5.

3.2 Theoretical framework

There is little present research on the group epistemic frames of faculty scientists. As described above, there is extensive literature on physics students employing multiple types of epistemic frames in the process of problem solving. We wish to utilize this study to apply the framework of epistemic frames to a new context. In this section, we will contextualize the theoretical framework of epistemic frames within the space of value-laden discussions present in the fellowship and the overall goals of this project. Doing so involves several different elements. First, we will briefly describe the characteristics used to describe the epistemic framing of group discussions, and then the assumptions of the epistemic frames tied to value-laden discussions.

3.2.1 Characterizing epistemic framing

In the moment of making sense of a present situation, the way an individual chooses what information and skills are relevant is framing[22]. Critically, this process of framing is dynamic, happening on the time scales of minutes to seconds. Instances of epistemic framing have been investigated in environments such as students working through problems together. This helped to highlight different ways that students considered problems, and this process allowed researchers to determine probable considerations the students made at that time.

Considering then the application of this process to the particular context of this work. Little research establishing epistemic framing in discussions among faculty has been done. If we consider that epistemic framing can be similarly studied for faculty discussions, then the process of identifying that framing experience may also be applied. Something notable is the difference in experience between students and faculty. As the faculty are considered to be experts, we can expect a difference in complexity in the framing experience[70]. Faculty scientists can be expected to process the present situation at a faster pace than the students. In the process of framing, students have been shown to activate particular resources. For faculty, being experts, we would presume that they would have a larger network of different resources to potentially work with. Of particular interest among these resources are ethical and epistemic values.

3.2.2 Epistemic frames of value-laden discussions

The collection of the information that someone puts together for a particular situation is the epistemic frame they form at that time. As someone tacitly places their understanding and perceptions in the context of the situation, they select the necessary sets of skills and information. The same must be true for groups of individuals in the same context. The group epistemic frame is something that likely elicits a particular set of resources from each individual as they share the group's making sense of 'what it is that's going on here'. An added complex layer to the frame is added by the interactions of the group. An epistemic frame is characterized by the expected behaviors and characteristic markers of the individual's making sense of the situation. A group epistemic frame adds an element of sharing and reacting to those characteristic markers. There must then be some relationship between the characteristic markers at the individual and group levels. For example, in a discussion between faculty scientists, if someone mentions a personal anecdote from their discipline related to the work they conduct, and someone else from a different discipline responds with another perspective, then the characteristic marker is a personal reflection coming from the same point of relating the discussion to the scientist's particular discipline. This is a supportive type of collective characteristic marker, and other relationships are likely present in the group epistemic frame. Another element of this set of shared characteristics is the discussion-laden nature of the group frame. As the group navigates the discussion, there will likely be instances of both collective understanding and arguing.

Capturing group epistemic frames is difficult. The collective frame cannot be directly observed, as described above. Instead, we must employ characteristic markers to determine the likelihood of an epistemic frame to exist. Extensive work has already been completed in education research on different characteristics that likely elicit the epistemic frames observed in both individuals and groups. Characteristics such as verbal cues, non-verbal behaviors, and warrants have helped to determine a number of existing frames in education research. Students working through physics problems revealed several groups of epistemic frames. The 4 frames determined by Scherr and Hammer helped to promote the use of characteristic markers such as behaviors and verbal cues[24]. The work conducted by Irving introduced the concept of frame coordinate planes with students engaging with problems through Math and Physics understandings, with additional work by Chari et al. to determine how students shift between these frames[23;25].

3.3 Methodology

We first establish the process of identifying group epistemic frames within the specific context of faculty value-laden discussion. Following established practice in the education research literature, the frames we establish are those shared by the group as a whole, rather than those held by individuals.

3.3.1 Context: fellowship formation

To explore how faculty engage in value-laden discussions on ethics and science, we formed a fellowship of fifteen faculty in the sciences at one Midwestern research university. Recruitment consisted of email invitations sent to department chairs in biology, biochemistry, chemistry, geology, and physics. The fellowship participants were split into two roughly equivalent groups to better facilitate small group discussions. These groups were then reconfigured halfway through the fellowship again to highlight potentially different interactions between the faculty.

The fellowship met ten times during the academic year. The topics discussed were chosen by the authors to help elicit a wide range of discussion possibilities, including both societal and scientific areas of interest. These topics included the social epistemology of science, scientific advocacy, and diversity. The faculty were also told that they would be explicitly addressing value-laden questions and that the overall research goals of the study would center on ethics in science. The sessions consisted of approximately 70 minutes of discussion, each organized as one member of the research group acting as moderator to each group of faculty engaging with the central topic. The faculty were given readings and prompts prior to each session concerning the central topic, and the moderator only intervened to provide additional context when necessary.

The first session introduced the faculty to philosophical concepts and ethical theories. The next six sessions of the fellowship were held in person, before they were moved to virtual discussion for the last three sessions due to the COVID-19 pandemic. The topics of the intervention were chosen to cover a range of relevant subjects at the intersection of science and society, including public participation and perception of science and advocacy. Subjects such as diversity, inductive risk, and statistical error were picked to encourage participants to reflect on their own understanding of the intersection of epistemic and ethical values. Crucially, this reflection also encouraged the faculty to relate the subjects to their own goals related to their research and work. With interdisciplinary participants, this would potentially allow for fruitful discussions of the general standards and goals of science, standards across disciplines, and the impact science has on society. Each of the fellowship sessions held in person was video and audio recorded from two perspectives.

3.3.2 Data collection

Seven of the fellowship sessions were held in person. We elected to consider only these sessions for this study as we presume a certain consistency in group interactions that would be disrupted when the sessions transitioned to the Zoom environment. This is allotted for approximately 490 minutes of discussion across 7 topics. The seven session topics investigated in this study are listed in Table 3.1. There are a total of 36 videos of data, including both groups of faculty. The average video length was just over 30 minutes.

The video and audio data were collected for each session. The structure of each session began with the moderator opening by introducing the general topic through the context of the papers provided. In some cases, this would additionally involve the moderator providing context for the present study. The time frame for this section often spanned only five minutes. The moderator would then open the discussion for the participants present, first centered on the general feelings and impressions of the papers involved. The participants were then responsible for the remaining

Session #	Topics
1	Introduction & Ethics
2	Value-Free Ideal and Inductive Risk
3	Value Choices & Statistics
4	Replication
5	Social Epistemology
6	Diversity
7	Advocacy

Table 3.1: *The first 7 sessions of the fellowship that were held in person. Each session lasted approximately 70 minutes, covering one topic. The moderator introduced the topic listed, and the session would consist of the participants engaging in different discussions surrounding the central topic and papers provided.*

time of the session to guide the discussion. The moderator would help to facilitate the discussion and provide context when needed for the study or the discussion itself. Finally, the moderator closed the session with any necessary logistical information for the next sessions.

3.3.3 Identifying group epistemic frames

Several authors gathered the necessary videos for the first seven sessions of the year-long intervention. Each session was approximately 70 minutes long, with some sessions in multiple videos. The process of identifying the epistemic frames happened in four major processes: identifying and characterizing shifts in framing, identifying segments of group interactions sitting between the boundary of each shift, gathering similar segments present from all sessions, and finally characterizing these groups of segments as group epistemic frames.

Shifts in framing

Coders watched each session multiple times. The first time watching the video allowed the coders to get a sense of the discussions present in each session. The second watch through was focused on identifying different characteristics of the group of participants. First, the number of participants was identified in each session. Each participant was given a letter as identification, starting with the Moderator as M, and proceeding clockwise around the table with A, B, etc., where participants would likely have a different letter for each session. From the literature of epistemic frames listed

Characteristic Markers	Description	Example
Gestures	Physical hand gestures by the participants	Participant A gestures to B while speaking
Tone of Voice	Vocal quality of the participants	Participant F speaks in a nervous voice.
Volume	Volume of the group discussion	High group Volume
Body Language	Distinct body language of a single participant	Relaxed, Shifting in Seat
Discussion Types	The theme of the discussion	Discipline Related
Claims	Verbal claims made by participants	Participant D stated mice research is more accessible
Values Invoked	Particular values invoked by a participant during discussion	Participant D invokes an Ethical Value about birth control accessibility
Number of Speakers	A number count of the number of participants speaking during one segment	6 participants
Change Instigator	Participant (including the moderator) responsible for changing discussion	Participant B shifts the discussion to statistical standards
Warrants	The evidence used by the participants in a discussion	Participant C mentions a personal research anecdote related
Atmosphere of Discourse	A check of the speaking atmosphere including how comfortable the participants appear	The discussion becomes tense and more argumentative

Table 3.2: *The characteristic markers identified from the Fellowship Session video data. This list was chosen using previous literature as a basis, with other markers identified during the process of watching videos. Each marker has a description and an example that is expected from the faculty sessions.*

above, we used this second time through the session to identify characteristic markers. Some markers, such as gestures, tone and volume in speech, warrants and types of discussions, and body language, were included as potential characteristics of the group frame. These characteristics were initially chosen based on the established markers used in previous frame studies. Other markers were included based on the nature of this study (values invoked), and others due to their presence in the fellowship sessions (change instigators and discussion atmosphere). These markers were identified as belonging to individual participants or the group itself. The first and second time viewing the fellowship sessions helped to expand the list of markers we considered. The description of each characteristic marker used and an example consistent with the value-laden discussions are all included in Table 3.2.

Watching each session again, we placed the Characteristic Markers at their appropriate time in the videos including the individual(s) that invoked each marker. This resulted in each of the session videos now having all characteristic markers now labeled.

Identifying segments of group discussions

The sessions were each around 70 minutes long. Based on the literature, we would expect the group epistemic frames to be on a shorter time scale compared to the entire video. We also observed multiple types of discussions happening in each session. Each session concerned one topic, but multiple discussions over different aspects of that topic occurred in each session. A Characteristic Marker we used in the coding of each session included the type of discussion the participants were having. Shifting the focus of identifying frames to identifying particular subtopics of discussions greatly reduced the time frame. Hereafter, we will refer to particular types of discussions as Segments of the session. These types of discussions ranged from participants reflecting on their own understanding of the paper introduced at the beginning of the session, to them sharing anecdotes of their personal experiences on a topic, to an argument of the particular session topic itself. A critical element of the framework is the 'making sense of the current situation', and in the event of discussions changing topics, we determine this as a suitable boundary between frames, as this is when the characteristic markers are likely to also change. Dividing the sessions into segments along the lines of changing subjects in each session was carried out by each coder. The coders additionally met several times to ensure that there was full agreement on the time frame of each segment, and that characteristic markers were correctly coded for each segment after the divisions. Any disagreement was discussed among coders until full agreement was reached.

Gathering similar segments from all sessions

With each session now divided into segments with characteristic markers placed, we first determined the number of segments per session. We then gathered segments based on several criteria. In an effort to remove biases on particular frames aligning with the different types of discussions of the fellowship, we primarily focused on the characteristic markers. Based on the literature above, this method of gathering together collections of characteristic markers is already well-established[24]. We investigated the segments in total, paying particular attention to groups of characteristic markers present in multiple segments. For example, segments where a particular set of behaviors and vocal cues are present during certain claims or values being invoked would be placed in a single group.

Characterizing group epistemic frames

Each group of segments now shares a list of similar characteristic markers. The characteristics of each segment are those present in the group of segments, and additionally, it notes how many of the participants show those characteristics. These groups share these characteristic markers, and by the metrics mentioned above, this denotes that these groups are the group epistemic frames present in this study. The characteristics of each group's epistemic frame are based on the characteristics present in these groups of segments.

The characteristics that describe the group epistemic frame are based on what the participants are doing (i.e., behaviors, vocal cues, voice tone) and how the participants are discussing the topic of that frame (i.e., the warrants, values invoked). In each group of segments, there are warrants that the participants invoke during the discussion. If we consider the warrants and values invoked in each discussion, we can infer a set of expectations that someone might have while in this epistemic frame. Consider what epistemic frame a Teaching Assistant working through a physics problem on a board while lecturing is in. The TA exhibits behaviors like writing on the board, talking with gestures as they lecture, and looking around the room for potential questions. The warrants invoked could include particular physics theories for the problem or personal interpretations. Considering both the way the TA is talking and the way that they are describing the problem helps us determine particular expectations the TA may have during that time. These expectations in this example may include how much time the TA needs to spend on a particular aspect of the problem or what information is critical to add to the current explanation.

Table 3.3: *The 3 group epistemic frames identified from value-laden discussions in the fellowship sessions. Each epistemic frame is described by the framing that the group collectively experiences, identified by the expected behaviors, warrants, and expectations they likely have while in that frame.*

Group Epistemic Frame	Expected Behaviors	Expected Warrants	Likely Expectations
Facilitator Frame	Writing, Listening to Speaker, Reading	Small clarification questions	Listening to a primary speaker and not need to form input
Sharing Science Frame	Normal Talking Rhythm, Gesturing, Leaning Forward	Experiences through career	Learning from one another.
Arguing Claims Frame	Elevated Pitch and Noise, Increased Gesturing, and Shifting in Seats	Experiences through personal values	Justifying claims

3.4 Results

Investigating the first seven sessions of the fellowship, distinct segments were identified and grouped into 3 distinct epistemic frames. The segments lasted a range of different times, from under a minute to the upper range of around twenty minutes. The average length of time for these segments was about seven minutes. The average number of segments per session was around eleven. The group epistemic frames are: Facilitator, Sharing Science, and Arguing Claims. The characteristics and potential expectations and warrants of the participants engaging in each frame are listed in Table 3.3.

For each of the following group frames, we will describe both the common framing experience that the participants engage in while in this frame and the characteristic markers identified by that framing experience.

3.4.1 Facilitator Epistemic Frame

Group framing experience

The group framing experience of this epistemic frame centers on the participants listening to one participant or the moderator for an extended period of time. The participants in this group frame have the least amount of group engagement in the whole of the fellowship discussions. The participants assume a role reminiscent of a student listening to a lecture. The moderator likewise assumes the role of the lecturer as they introduce new concepts and ideas or elaborate on points being made.

Characteristic markers of the group frame

The characteristic markers consistent with this group epistemic frame include instances of faculty writing down information, reading through documents, and primarily listening to the speaker, who is often the moderator. The faculty may raise questions or comments based on the context of the situation, most often seeking clarification for the discussion. The expected warrants made during these discussions show the participants using this time to engage with the material being presented and ask questions when something was not clear to them.

3.4.2 Sharing Science Epistemic Frame

Group framing experience

This frame describes a group of faculty scientists engaging in group discussions towards a shared understanding of a concept. This manifested during discussions where the participants divided into groups depending on their familiarity with the central topic. The topics that were discussed tended to be interpreted by the participants concerning their own experiences in their careers, with the following discourse focused on the way that the participants are reaching each other's views on the subject. The following transcript was taken from the Replication session to show this type of framing experience.

E: ...what have we missed by always doing mice for cancer?

D: I mean I agree with you, but on the other hand you can see the utility right? ....and you do your test, and so there are less variables to worry about and you come up with a result. ...Let's see if we see this in a more complicated system; chimpanzees or whatever (**hand gestures**). But, but at least with the mice, you get rid of a lot of variables. (**epistemic value**)

B: Absolutely. I'm not saying mice studies aren't worth-while. (**epistemic value**)

E: But, where I come into agreement with if you draw all your conclusions with the mice and make the jump to humans that's where I see you are expressing, I think, the problem.

B: I mean mice, 100 percent have a role within human biology studies. But there is a reason they are pre-clinical and not small clinical studies you got to have a lot mouse evidence before you go to a human because I'm not taking your drug because two days ago you cooked it up in your laboratory and decided it worked...you better have evidence before you move onto the next level. (**epistemic value**)

F: On the same lines right, living beings are so complex the positive controls they use are different estrogen mimics right obviously.(**muttered agreement from several participants**). So the response in different species can be different on very narrow bases between two different chemicals. And so it's tricky, it's real tricky I think to conclude anything really positively about the effect humans from studies on mice from different positive controls...I think the most interesting thing I found is the most simplest thing and that is they report that 100 percent of industry studies found no effect (**laughter and mutters of agreement from several participants**)

G: I think half the studies showed no effect, right so half of them are industry studies...

C: It was hard to get through, but I thought there was an undercurrent of the author suggesting perhaps that firms that have undertaking this research for profit. I thought they were coming out and saying there was an implicit sort of thing they were not reporting

G: I'm sort of amazed how, I'll just say, political this paper is (**some participants laugh**). I cannot imagine a physics paper like this taking shots at the FDA and companies like the EPA....from my field, ours is so dry you would never make any sort of reference.

Characteristic markers of the group frame

The characteristic markers that help to define this group frame include the behaviors and verbal cues of the participants, such as the comfortable laughter and agreement, both verbally and physically expressed in the example above. To contribute to the collaborative and comfortable nature of this group frame, we found the participants having a tendency to gesture more as they acknowledge contributors to the discussion and lean forward while visually keeping track as the discussion passes between speakers. Speakers referencing particular points of the paper being discussed led to other participants finding the point in the paper to follow along.

The expected warrants that the participants connected to the discussion tended to follow a particular pattern. The discussion tended to center on more scientifically-minded topics, such as cancer research for mice in the transcript example above. During these discussions, some participants would be able to speak from a discipline-specific set of experiences and knowledge to discuss the topic. The participants in these segments would likewise invoke values that are consistent with these warrants. In the language of the categories of values determined by Linville et al.[?], the values invoked in this group frame are in the epistemic category, focusing on what the participants consider advancing or improving science. The warrants invoked include how some participants are able to relate to the topic with this knowledge and experience. For the participants of the discussion less familiar, they can use this opportunity to ask clarifying questions of the more expert participants or quietly listen. This group frame also exhibited high engagement from the participants, with many segments having all participants speak in the time frame of a single segment. In total, the participants in this group frame use this opportunity to learn from one another about a particular topic that is specific to a discipline that a participant may have expertise in.

3.4.3 Arguing Claims Epistemic Frame

Group framing experience

This group frame consists of faculty scientists engaging in group discussions where statements and claims made by the participants are being challenged at the level of values being invoked. This was marked by there additionally being a shift in the atmosphere in the room and discussion during these segments. The faculty acknowledges this change in the speaking atmosphere, and the response to this change invokes particular characteristic markers distinct from the group frames identified above. The following example is a transcript taken from the Diversity session to showcase this framing experience.

> M: So what about the policies and should we be stepping into that, so what do you think about the actual policy recommendations?
>
> *Pause for 5 seconds*
>
> B: I mean I am all about making birth control accessible but forcing it on people would be incredibly difficult... (**ethical value**)
>
> E: While we're talking about populations, they are pretty low on things I personally worry about, there's like lots of evidence that as cultures reach this point where they're reasonably prosperous and women are reasonably empowered and have access to reasonable contraceptive approaches ... I think it sort of takes care of itself in the long run just because of a (**F shakes head**) assuming we're able to build that sort of society but...equality where women are adequately empowered and have access to resources
>
> B: Which they're not, they're not (**ethical value**)
>
> E: No, but it's, it's happening around the world, not everywhere but you know
>
> B: It's hard to get birth control in some places in this country, it's expensive too right?
>
> F: Just like related to levels of education...(**cut off by C**)

C: So...that's much more complicated in that, right because it's related to how much energy you consume and you cannot consume energy if the right to consume it per capita in the US if you tried to do it all over the whole world...I mean the reason that people don't have like 500 kids here and in western Europe is because it's very very expensive and they have other things to do.. (**Other Value**)

E: I mean I don't know, I wouldn't have more than two kids even if you know even if quote, unquote, the whole experience. You know it's too much work

C: You cannot accept it like...(**overlapping with E**)

E: You know, it's like there's this huge crowd of people who wanna have 5 or 6 kids.

C: There's a lot of people in Kansas that have 4 or 5 kids right?

B: They totally exist...they exist (**overlapping with C**)

F: There's a lot of evidence that related to levels of education and you can have the educational system. There's statistics there that if you wait to have a kid until you got your undergraduate degree then the world population would actually go down.

C: Yeah, I mean but that's only if people actually get an undergraduate degree right?

F: Right right but that's as an example (**gesturing more**). Basically every, every piece of knowledge helps to kinda...

C: Yeah but everyone with an undergraduate degree uses much much more energy than people who work in fields after dropping out of sixth grade... I mean I don't think there's... I mean we don't have that kind of energy resources to do that with seven billion people...there's just no way to deal with global warming... (**Other Value**)

Both D and F start talking

D: Right besides that....(**gestures for F to continue**)

F: but just as a, just....increasing knowledge basically of you know...how to.... (**gestures back to D**)

D: So statistically...I agree with ‘C’ that well ok I mean but not everyone is going to have these undergraduate degrees, but on the other hand something else that was mentioned in our reading (**gestures towards own computer**) is that more and more women have job opportunities and it seems to be heading in the direction it seems of equal pay for equal work... (**Other Value**)

Characteristic markers of the group frame

The markers that help to characterize this frame are representative of a change in the discussion atmosphere felt by the participants. The behaviors are reflective of this, with faculty noticeably shifting positions in their seats, gesturing more, and avoiding eye contact. The tone of voice tended towards nervous, with participants getting caught up in phrases. More significant pauses take place between points in the discussion, with participants appearing less confident in their speech.

Considering then the arguments being made in this group frame, with particular attention to the warrants invoked by the participants. The discussions of this group tended towards personal anecdotes of the participants. This would often take the form of discussions starting from discussion topics that participants relate to their field expertise, to more personal anecdotes. The values invoked in this group frame are reflective of this. The category of values invoked in this frame follows the ethical category as described in Section 2.1.3. These values are more personal for the participants, considering their background with the topic. The warrants of the participants in this group frame reflect the values invoked by the participants. The way participants shifted towards personal anecdotes to justify arguing for or against the claims made during the discussion.

3.5 Discussion

3.5.1 Group epistemic frames of value-laden discussions

The methodology described in Section 3.3 was informed from established literature[24;25;84]. We presumed in this study that a methodology shown to identify epistemic frames of group interactions for students would also identify frames for faculty. The notable difference between the student interaction and the faculty interaction is the level of expertise. In epistemic frames and framing literature, this difference is quantified in epistemic complexity and ultimately shows in terms of how quickly experts navigate through an activity[70]. The fellowship sessions consisted of value-laden discussions, so elements of these discussions were added as markers used to identify group frames. In particular, identifying the values invoked by the participants and the warrants invoked provided additional opportunities for understanding the way that the participants were making sense of the present discussion. Considering each of these markers as individual epistemic resources[22], we then identify the resources that are activated, or suppressed, by the participants as they collectively make sense of the discussion. The way this network of epistemic resources is activated by the collective participants is consistent with what has already been done in previous studies[80]. The collective 'making sense' of the present activity identified in previous studies is not a methodology intrinsic to those particular contexts. The group frames identified in this study also exist within defined boundaries, exhibiting particular characteristic markers that include the warrants and values invoked by the participants. With this in mind, we conclude that this adapted methodology is suitable for identifying group epistemic frames for faculty discussions.

3.5.2 Group epistemic frames consistent with known frames

The group frames in this study were identified in the context of faculty value-laden discussions. Other epistemic frames and framing experiences have been applied to multiple contexts. For example, the frames identified by Chari et al. (i.e., Conceptual Physics, Algorithmic Physics, Conceptual Math, and Algorithmic Math) were used in the context of student and instructor understanding of upper-division physics problem-solving[23]. Modir et al. then considered how students navigate these 4 frames, particularly in upper-division quantum mechanics courses[84]. As frames transcend the context they were first identified, we then wish to examine significant similarities and differences with the established frames in Physics Education Research. For many of the frames and framing experiences, the named frames often follow a naming convention consistent with those particular contexts, such as the Conceptual Math Frame from Chari et al.[23] and the Physical Mapping Framing Experience of Bing & Reddish[22]. The group frames identified in this study are heavily based on group faculty discussions, so the comparisons made to past frames and framing experiences include characteristic markers and the general framing experience itself.

The Facilitator Frame is similar to what we might expect from a lecture-style situation. A participant in the discussion takes the central speaking role for an extended time (i.e., on the time scale of minutes), and the rest of the participants recognize that the speaker has a certain expertise in the discussion. This can be seen as the non-speaker participants ask questions of the speaker or make notes of what the speaker is saying. Most often, this group frame centers around the Moderator as they introduce a particular topic. This framing experience, and the situation as a whole, is reminiscent of students engaging with lecturers or Teaching Assistants. A notable comparison is in the *TA Frame* identified by Scherr & Hammer. In this frame, the behaviors of the students include sitting up straight, making eye contact with the TA, and reducing gestures. The expectations of the students in that frame include that attention belongs on the TA and rehashing thinking[24].

The Sharing Science Frame is a group discussion frame with high levels of participant engagement. In the fellowship discussions, instances of particular subject matter naturally identified participants with more experience and those with less. These discussions evolved into instances of those ‘more expert’ participants highlighting any personal insights connected to the present topic. The response of other participants took several paths, with some remarking with surprise on the way particular disciplines considered the topic or relating to their anecdote, potentially from a different discipline. This collaborative framing experience is similar in nature to the *Discussion frame* also identified in the Scherr & Hammer study. In this frame, students gestured more, were more animated, and made more eye contact. They expected that peers would be watching and want to understand the present activity, sustained emotional and/or intellectual engagement, and attention belonging to peers[24]. We find similar characteristic markers in the group frame of this study, where the participants were leaning forward, actively following the discussion as it changed speakers, and speakers gesturing to other participants to acknowledge what had been said. The expectations of the participants in these fellowship sessions focused on learning from one another and sharing their experiences.

The Arguing Claims Frame is another discussion-focused frame, but one that is more difficult to directly compare to the established frames and framing literature. In the *Joking frame* from the works of Scherr & Hammer, behaviors of the students included fidgets and unsettled gazes, and the expectations included embarrassment and a perceived vulnerability[24]. Critically, the students experiencing this vulnerability are often supported by the other students in the group to reassure and support them. In the Arguing Claims Frame of this work, we identify similar characteristic markers of faculty appearing uncomfortable, shifting in their seats, or appearing to disengage from the discussion. A noted difference here is where the support of each faculty comes from, where faculty appear to choose sides of the arguments. The level of engagement in this group frame is also mixed, with some segments seeing higher engagement levels with many faculty speaking and engaging the discussion, or times when nearly half of the faculty engage.

3.5.3 Group frames productive for ethical interventions

Considering the engagement levels of the faculty discussions, we identify 3 distinct levels: low (1-2 faculty), medium (3-4 faculty), and high engagement (5-7 faculty). The Facilitator Frame is characterized by low group engagement, where one participant is primarily leading the discussion by themselves. The Arguing Claims Frame has mixed engagement, depending on the context of the discussion, and the Sharing Science Frame with high engagement, where all participants are often engaged in the discussion. Part of the motivation for the Goals and Values in Science project is to consider the impact that value-laden discussions have on someone's ethical decision-making. Another study of this project established that value-laden discussions have an impact on ethical decision-making[21]. Two of the group frames identified in this study are partially characterized by the values invoked by the participants in each segment. In particular, the Arguing Claims Frame considers moments when the participants invoke more ethical values. An avenue of further analysis considering this work is how participants shift into group frames. If we presume that group epistemic frames where participants invoke ethical values have an impact on the ethical decision-making of the participants of the intervention, then understanding how to promote faculty to engage in these epistemic frames becomes the next step of this work.

3.5.4 Disjointed and mixed group frames

The segments found in this study were grouped by characteristic markers to determine the group frames. The three group frames discussed in Section 3.4 were determined using the majority of segments found, but not all segments were used to determine these group frames. A critical element of identifying epistemic frames for group activities is the basis that stability of the group frame is built on the participants contributing to a single network of activated epistemic resources. Where the participants share an intuition towards making sense of the situation at nearly the same time[24;73;80]. The three group frames all share characteristic markers that show strong evidence for the epistemic resources being activated in a coherent network during these framing experiences.

We additionally identified segments that do not exhibit this type of cohesion in resource activation. An example of this could be a segment of fellowship discussion where half of the participants in the room are in a Sharing Science Frame and the other half of the participants are in an Arguing Claims Frame. The characteristic markers in this segment reflect these two frames, with characteristic markers of the segment reflecting both group frames. Segments like these present interesting opportunities for further analysis into the ways that faculty engage in group discussions when not seemingly aligned in goals.

3.5.5 Study limitations

The information found from the study provides a starting opportunity for understanding group frames. We identified discrete segments that lasted for an average of seven minutes each. These time frames are in line with the established literature. However, this does not mean that different frames could exist within the scope of different identifying criteria not described above. The characteristic markers used are built from the ways that the participants behaved and argued through the discussions as a group, and these markers could be different for other types of value-laden discussions.

The participants involved in this study also do not represent all types of faculty scientists. We had a relatively diverse set of disciplines within the fellowship, but having very few numbers of each discipline. While this is acceptable for generating preliminary hypotheses regarding the epistemic frames faculty use in ethics-focused discussions, a robust understanding would require a larger number of participants across different institutions and career stages to get a better understanding of how each discipline regards the different discussions offered by the sessions.

Chapter 4

Investigating the structures of physics graduate education

4.1 Motivation and outline of this chapter

The concerns of graduate education in physics are well-established, some dating back to the early 20th century. Investigations of student difficulties and institutional barriers have promoted promising advances in understanding the varied experiences of graduate students. A more holistic review of graduate education has been done nationally for STEM disciplines, but one has not been specifically done for physics. In this study, we are interested in beginning this holistic review starting with the explicit requirements and practices described by the physics departments across the country through a lens of structural analysis. The research questions to help guide this investigation are:

- What are the common structures of physics graduate education?
- How can graduate physics programs be informed by identifying these structures?

To accomplish this, we center the analysis of Structures as the vehicle of investigating the different practices in physics graduate education in Section 4.2. Section 4.3 discusses the process of gathering and analyzing public documents used for this study. The identified structures are described in Section 4.4, each with the different requirements as described by the physics graduate departments themselves. Section 4.5 will then consider the impact these structures have on physics graduate education in concert with the established literature.

4.2 Structures of physics graduate education

The focus of this study centers a structure-deficit model of physics graduate education. The studies listed above highlight long-standing concerns with graduate education such as candidacy exams and the dissertation. An element of these concerns includes the effect that these concerns have on both the graduate students and the system of graduate education itself such as student retention. We propose then, that there exists an environment of graduate education with elements in this environment such as candidacy exams. In this way, we identify an environment just as Cotner & Ballen determined in the classroom and Webb & Paul determined for a physics course[105;107]. In this environment of physics graduate education, we additionally presume that the elements of this environment have a significant impact of the participants. For example, an established and known element of graduate education is candidacy exams. This element has shown to have significant impact on student experiences and retention. In the model of structure-deficit, the candidacy exams is an element of the physics graduate environment that impacts the students. We then identify these elements of physics graduate education as individual structures of the environment. In a structure-deficit model, these individual structures each impact a student's experience.

4.2.1 Source of structures

The structures of a physics PhD for this study involves two systemic parts: the graduate college of the institution that offers a graduate degree, and the physics program of that same institution. The graduate college is often a separate entity with its own regulations and guidelines. The graduate college must establish regulations for all necessary departments and determine what requirements are necessary for all graduate students of all relevant disciplines. Those requirements and expectations are then disseminated to the individual programs. We then consider the department, physics in our case, as the other major entity in this study. The physics department has primary control over monitoring the progress of its graduate students. The program itself serves as the primary provider for graduate requirements for its students, with the institution's graduate college serving supplementary requirements when needed. The interaction of program requirements with the regulations of the associated graduate college is the full extent of the available information for this study.

4.3 Methods

4.3.1 Data collection

To understand the diverse graduate program requirements, we first need to determine what the programs across the country consider requirements for their graduate students. The American Institute of Physics (AIP) has collected information from physics programs in the United States for decades. This information has historically been collected and self-published each year, with most current forms of this data hosted on the AIP-owned website GradSchoolShopper (https://gradschoolshopper.com/). GradSchoolShopper additionally contains a list of the programs across the country that grant graduate degrees in physics. The authors accessed and collected the list of programs in the spring of 2023. This list includes all PhD and Master's programs, including interdisciplinary degrees such as biophysics, medical physics, etc. The total number of programs identified in this list was 275. To identify program requirements for the Physics PhD, we limited the full dataset to include only those institutions. This brought the number of institutions in the dataset to 190. In the interest of determining diversity in the results of this dataset, we elected to additionally constrain the list, excluding those programs listing a subject other than solely Physics. This includes programs named Applied Physics, Biophysics, Astrophysics, etc. The remaining programs available for analysis in this study to 181 PhD-granting institutions.

To efficiently collect data from each institution, we used publicly available data. Utilizing public data allowed for easier and faster access rather than interviewing multiple faculty at each institution, and uses the data most readily available to prospective and current students. We prioritized the use of handbooks released by each department. When the handbook was not present or provided insufficient information on requirements, we supplemented information from the website of the department. We further sought the graduate handbook of the institution's graduate college when information was limited in the physics department, or directly referenced the graduate college for more information. We collected information from each data source based on the definition of program requirements outlined above.

4.3.2 Document and thematic analysis

A process of Document Analysis and Thematic Analysis was adopted to read, analyze content, and interpret the information present in the data[88;119;120]. We expected a number of categories to be present in this study. The work completed by Donkor & Harshman[88] established a set of known categories present in chemistry graduate education, alongside the present understanding of the authors' participation in graduate programs helped to develop a set of categories. We began with a pilot group of five programs, holding extensive discussions between the authors on the data as we read and cataloged it into the established categories. We additionally accounted for different practices present in the data by adjusting the groups of categories and adding subcategories when needed. For example, the inclusion of core courses used as exemptions to reach candidacy was due to their presence in the data. In this way, we adjusted the set of categories to account for these emergent practices when they were persistent in the data. To ensure consistency in the data, different practices were marked in the data whenever encountered and discussed among the authors to determine if the practice is a suitable requirement for the study based on the definition listed above.

The resulting categories used for the remainder of this study are based on the full data set. The requirements fall into 10 structures, including: Core Courses, Elective Courses, The Comprehensive Content Exam, The Research Readiness Exam, The Dissertation, Supervisory Committee, Annual Reports, Training Modules and Courses, Teaching Assistant Requirement, and Presentations of Research.

4.3.3 Calculating inter-rater reliability

The coding process was shared between 5 individual coders. At the time of introducing another coder, we taught the coding process through inter-rater reliability (IRR). Three authors chose 12 institutions to code. Fleiss Kappa was used to determine the rater agreement during the initial discussion of the codes for each institution[121]. The metric determined between the three coders was measured as 0.652. Landis & Koch determined that this level of agreement is substantial[122]. Discussions were held to ensure 100% agreement among coders before the final codes were entered into the data sheet for this study.

4.3.4 Data limitations

The results of this study are based on public data. We prioritized graduate handbooks for each university, but necessarily used the websites if the handbooks were not available or proved insufficient for data collection. By using this type of data, we acknowledge that what is listed in the public data may be inconsistent with what is actually practiced at each university. We further note that while true, any inconsistency may be difficult for aspiring graduate students attempting to find credible information from a prospective program.

The coding for this study took place over a calendar year, with several authors acting as coders for the dataset. Several instances of teaching the coding process took place, with many discussions during that time frame to work through difficulties or questions about institutional language and codes. A significant portion of the coding was done individually, with one person responsible for one institution at a time. With internal efforts to ensure that discussions helped preserve the integrity and validity of the data being coded, there exists a possibility that the codes may be incorrect.

4.4 Results

This section will describe both the common structures identified by the analysis of 181 physics graduate programs and the requirements of each structure. We will define each structure and requirements, and list the number of institutions with explicit mention of each requirement. At this time, we will again highlight that the number of institutions in this section is those that had language of the requirement. Each section will additionally highlight notable elements of each structure.

4.4.1 Courses

Courses involve two structures: Core Courses and Elective Courses. Each structure has a number of requirements as outlined above in Tables 4.1 and 4.3. A common requirement that concerns both structures is the Grade Point Average, hereafter G.P.A. 142 institutions require students to maintain a certain G.P.A., with 129 institutions requiring a 3.0. We identify little variation in this number, with two institutions requiring a G.P.A. under 3.0, eight institutions requiring a G.P.A. over 3.0.

Core Courses

The Core Courses structure identifies requirements surrounding the courses required by all graduate students in each program. These courses were identified by the programs as 'Core Courses', or named as courses that every graduate student must take. The Core Courses structure has 8 requirements. The number of institutions that have these requirements is shown in Table 4.1. First, we consider the total number of courses required. We determined this by summing the list of courses from each program. 168 programs specifically list the courses required by students, with 114 programs also specifying how many credits each course has. The range of total core courses required was between 0 and 12, with an average of total core courses required being just over 7. Notably, the remaining programs provide alternative ways for students to take courses, such as students working with their advisors on which courses to take. The G.P.A. of a program is a common, known requirement in education at all levels. 6 programs in this data require students to maintain certain grades related only to the core courses. If students are expected to finish all of their core courses by a particular year or semester/quarter, 9 programs have a Timeline requirement. The core courses are often required by name, with 173 programs requiring specific subjects for core courses such as

Requirement	#	Description
Total Number	168	The total amount of core courses required
Credit Hours	114	Number of core course credit hours required
Grade Requirements	6	Grade requirements specifically for core courses
Timeline	9	Expected semester or year when all core courses should be complete
Subjects	173	The required subjects of the core courses when named
Substitutions for Sub-Discipline Courses	16	Courses offered that follow specific sub-disciplines in the program in place of the required course
Options from List	32	Students allowed to choose a subset of available courses
Course Outside Department	9	Option for students to take courses outside the physics department

Table 4.1: *The requirements of the Core Courses structure. Each structure includes the requirements and a brief description of the requirement determined during the coding process.*

Quantum Mechanics or Statistical Mechanics. We identified some notable breadth in the subjects of core courses. Table 4.2 shows the most common subjects, with Graduate Seminar being the last shown in the table as the number of institutions requiring subjects after this is required by 10 or fewer institutions. Another element of subjects involves how many times the subject is required. For example, if an institution requires Quantum Mechanics I and II then we would count Quantum Mechanics as a subject that is required twice. Table 4.2 also shows the number of times each subject is required.

While programs list a particular number of subjects students need to take, there are some notable exceptions. 16 programs also allowed students to switch a common core course for an alternative that is more aligned with their particular research interests. For example, students may be able to substitute Quantum Mechanics 2 for another course like Solid State Physics. 32 programs allowed students to choose from a particular list of core courses, and 9 programs allowed students to choose a course outside the department to satisfy a core course requirement.

From this structure, we identify some notable findings. First, from the common subjects, we find that there appears to be a consensus for what physics graduate programs consider to be the core courses. Classical Mechanics, Electrodynamics, Statistical Mechanics, and Quantum Mechanics are required by more than 70% of the 173 programs that specify the subjects graduate

Table 4.2: *Most common subjects for Core Courses. The total number of programs that require that subject (T) is given, as well as a breakdown for the number of semesters/quarters (N). For example, 39 program required one semester/quarter of Quantum Mechanics, and 117 program required two semesters/quarters.*

Subjects	N = 1	N = 2	N = 3+	T
Quantum Mechanics	39	117	5	161
Statistical Mechanics	147	8	0	155
Electrodynamics	74	73	2	149
Classical Mechanics	121	6	0	127
Math Methods	54	15	0	69
Graduate Seminar	18	6	11	35
Other	112	7	3	122

students must take. Alternatively, outside these four common subjects, there is little consensus on the other courses that students should be taking in their programs as core courses. We identified 41 distinct subjects of core courses in this study, with many falling in the 'Other' subject in Table 4.2. Additionally, 117 programs require 2 instances (i.e., 2 semesters or 2 quarters) of Quantum Mechanics, significantly more than the 39 programs that require 1 instance, or the 5 programs that require 3. Electrodynamics, alternatively, sees the number of times the course is required nearly split between 1 and 2 instances. With this in mind, we find that there is a consensus on how much Quantum Mechanics students should be taking, but no consensus on how much Electrodynamics students should be taking. At this point, it is not clear how much the programs that operate on Quarter timescales are affecting this. 18 programs in our data are operating on Quarter timescales, and the amount of influence these programs are having on this data is subject to further analysis.

Elective Courses

The Elective Courses structure identifies the requirements of courses after the core courses. These courses are often described as an opportunity for students to expand the breadth of their physics content knowledge. The 8 requirements of the 129 programs that mention the elective course are listed in Table 4.3 and reflect this ambiguity. These courses are not listed by name, and are not courses that every graduate student in the program is taking. Instead, the language surrounding the elective course structure considers a number of courses or credits that students are expected to take.

We see a near-even split in the number of programs that list the number of elective courses required for the degree at 63, and the number of elective credits required at 64 programs. To determine the potential variable number of courses required by the programs, we needed to convert the number of credits required to the number of courses. Many institutions specify the number of credits a course is associated with. We used the number of credits per course as a way to convert the number of credits required to the number of courses required. Doing so reveals that the range of the number of elective courses required is between 1 and 13, with the average number required at just under 4.

Some requirements are similar to the Core Course structure. 2 programs in the data specify that students need to maintain a particular G.P.A. in specifically the elective courses, and 4 programs describe when the electives should be finished. The remaining requirements reflect the nature of the structure, where students have more freedom in choosing the courses that expand their breadth of physics content knowledge. 24 programs provide students with different courses aligned with particular physics sub-disciplines, such as astrophysics or high energy physics where students can choose classes. 6 programs also allow for students to partially satisfy this structure through courses outside the department. In keeping with the motivation to have students expand their breadth of content knowledge, 9 programs require students to choose elective courses that are aligned with the students' chosen research sub-discipline, whereas 8 programs require students to choose courses that are not aligned with those research interests.

4.4.2 Candidacy Exams

The process of reaching PhD candidacy is one all graduate students go through. We have identified two structures that impact the process of reaching candidacy: the Comprehensive Content Exams and the Research Readiness Exam. Several requirements span both structures. First is the grading element of candidacy exams. At the time of this study, we identified grading associated with students reaching candidacy, but not particular grading schemes used for either structure. Similarly, we identified when students were expected to finish all candidacy exams. 116 institutions have language to inform students when they expect candidacy exams to be taken. The process of PhD Candidacy often involves multiple forms of exams for many institutions. Each structure has 2

Requirement	#	Description
Total Number	63	The number of courses required. Does not include research credits
Credit Hours	64	Number of course credit hours required
Grade Requirements	2	Grade requirements for courses
Timeline	4	When courses should be taken
Options from List	24	Students allowed to choose a subset of available courses
Course Outside Department	6	Option for students to take courses outside the physics department
Course Aligned with Research	9	Students to choose courses aligned with the research field
Course Not Aligned with Research	8	Students to choose courses not aligned with research field

Table 4.3: *The requirements of the Elective Courses structure. Each structure includes the requirements and a brief description of the requirement determined during the coding process.*

Table 4.4: *The two structures of Candidacy Exams are Comprehensive Content Exams and Research Readiness Exams. Each structure has 3 formats: written, verbal, or both. We identified 10 distinct ways that programs use different formats of each Structure listed in this table. For example, 7 programs require students to take both formats of the Comprehensive Content Exam and only the verbal format of the Research Readiness Exam.*

		Research Readiness Exam			
		NR	Written	Verbal	Both
	NR		0	21	35
Comprehensive	Written	9	0	53	36
Content Exam	Verbal	0	0	4	5
	Both	4	0	7	5

formats, written and/or verbal. We identified a variety of combinations in the ways that institutions require these different structures. For example, an institution may require a written Comprehensive Content Exam and a verbal and written Research Readiness Exam. This is one combination of the 15 possible that can be made with each version of the structures available. The number of institutions utilizing each combination can be found in Table 4.4.

Requirement	#	Description
Subjects	86	Subjects covered in exam
Number	85	The number of exams required in total
Written Exams	114	Written Exams
Verbal Exams	25	Verbal Exam
Timeline	116	When students should take all exams
Grading Scale	12	Grading scale used for exams
Number of Attempts	84	How many times the student can take the exam
Exemption through Courses	18	Option for student to bypass candidacy requirement through a course
Early Attempt	5	Option for student to complete a requirement for candidacy at start in graduate career

Table 4.5: *The requirements of the Comprehensive Content Exam structure. Each structure includes the requirements and a brief description of the requirement determined during the coding process.*

Comprehensive Content Exam

The Comprehensive Content Exam is the structure that determines the requirements of Candidacy Exams that assess a student's knowledge of physics content. This assessment can be done in the form of a written and/or verbal exam. This structure has 9 requirements. The number of institutions with these requirements is listed in Table 4.5. Considering first the two formats, 114 programs require a written form of this candidacy exam. 85 programs specify the number of exams that students will be taking, and 86 programs specify the particular subjects. The subjects mentioned are listed in Table 4.6. We combined subjects with similar-sounding names, such as 'Particle Physics', 'Atomic and Particle Physics', and 'Nuclear and Particle Physics', as the particular subtleties of each subject were not captured in this study. Other instances of combining subjects occurred between 'Electrodynamics' and 'Electromagnetic Theory', and also between 'Classical Mechanics' and 'Advanced Dynamics', as this was something done in the process of coding the core courses. 14 programs also defined that the level of the subjects in the exams would be at 'upper undergraduate to lower graduate' levels of physics content knowledge, and 12 programs defined the grading scale that would be used for the exams. 25 programs required a verbal form of this exam structure. The general description of this exam type placed students verbally working through problems on a board in front of a committee of faculty who would ask them questions that may involve the subjects in Table 4.6. Students were generally expected to advance through this type of candidacy by the first or second year, with 116 programs specifying when the student would be expected to pass the exams. 84 programs gave students a particular number of attempts to pass this structure, with 2 attempts being the most common, in 65 programs. 5 programs allowed students a chance to take the exams early in their career with no penalty to them if they did not pass. 18 programs allow students to bypass this candidacy structure through the student passing particular core courses, where the subject of the core course aligns with the subject of the exam.

Table 4.6: *The subjects of the comprehensive physics content exam mentioned in the data. The subjects required in this structure were mentioned by 86 programs.*

Subject	N
Electrodynamics	86
Quantum Mechanics	86
Classical Mechanics	82
Statistical Mechanics	76
Special Relativity	13
Atomic / Nuclear / Particle Physics	13
Mathematical Physics	9
Modern Physics	8
Thermodynamics	5
Optics	2
Solid State Physics	2
Computational Physics	1
Astrophysics	1
General Physics	1

From this structure, we identify a few notable findings. First, the primary mode of assessing the physics content knowledge of the structure is done through a written examination, with the verbal form of the exam being uncommon at only 25 programs. Similar to the findings of the Core Course structure, the consensus of assessing a student's content knowledge is centered on the subjects of Classical Mechanics, Quantum Mechanics, Classical Mechanics, and Statistical Mechanics. Something interesting is how some programs describe that graduate students should have a particular level of physics content knowledge, near the level of 'upper-division undergraduate to lower graduate' of these courses.

Research Readiness Exam

The second structure of candidacy exams is the Research Readiness Exam. This structure assesses the way that a student is conceptualizing the research area they are interested in. This structure can also be done in a written and/or verbal format. This structure has 6 requirements. The number of institutions with these requirements is listed in Table 4.7. The common form of this exam is verbal, often given as a public talk in front of the Supervisory Committee, as described by the 166 programs that require it. The most common version of this candidacy exam is a thesis proposal or dissertation proposal. The student proposes the research they expect to conduct for the dissertation, based on their present understanding of the topic, any progress they have already made towards the dissertation, and where their work sits in the literature of that topic. We note here that while this is the common conception of this exam, the way it is discussed in the data varies significantly. We identified 17 different naming conventions for this structure, including Dissertation Proposal, Prospectus, The Qualifying Exam, and Diagnostic Evaluation, to name a few. The focus of the exam is similarly varied. While the student presenting work towards the dissertation is the most common focus, there are other forms, including a Literature Review of the topic of interest, a separate Research Analysis that is separate from the student's interest, or the student conducting a stand-alone research project determined by their committee. 59 programs describe the focus, although briefly, of the exam structure using language around these naming conventions. 81 programs additionally require some written element of this structure, where 31 programs establish a particular requirement for the written portion, which ranges between 5 pages to 20 pages. 116 programs specify a particular time frame that the students need to pass this exam structure by, often by the end of their second or third year. 12 programs additionally impose a grading scale for this structure.

Findings to highlight from this structure include the strong consensus that programs are interested in students presenting some amount of research a few years into their education. The subject matter or focus of the assessment leans towards students presenting a proposal of the work they will conduct for the dissertation, which is not as universal as the practice itself. There is an interesting variety in the operationalization of this structure, particularly between students presenting a

Requirement	#	Description
Written Exam	81	Written Exam
Verbal Exam	166	Verbal Exam
Timeline	116	When students should reach candidacy
Grading Scale	12	The grading requirement used
Focus	59	The purpose of the exam. This includes the subject
Length	31	Length of report or talk for the exam

Table 4.7: *The requirements of the Research Readiness Exam structure. Each structure includes the requirements and a brief description of the requirement determined during the coding process.*

dissertation proposal and presenting on a subject that is fully distant from their research interests. Further analysis to highlight the different ways that programs implement this type of candidacy exam may help highlight what programs value in this assessment. Another particular element of this structure is the intersection of the written and verbal types of this exam. In the case of the written form, it is always tied to the verbal format. There are no programs identified in this data that only require a written form of the structure.

4.4.3 The Dissertation

The Dissertation is the culmination of years of student research and progress towards becoming a subject-matter expert. It is the most recognizable structure in graduate education, yet it only has 3 requirements. The number of programs listing each requirement is shown in Table 4.8. All 181 programs in this data set contain language of the dissertation, but offer no further explanation of the purpose of the document and little information about how the writing of the dissertation is done. The little information found in this study on the dissertation refers to the practice as a reflection of the student's ability to do independent research. For many programs, the discussion of the dissertation in the data considers the forms necessary for the program's Graduate College. A requirement that ties to the dissertation concerns the 70 programs that list Research Credits tied to the dissertation by name. The only other requirement of this structure is the time frame that programs expect students to defend their dissertations by.

The most notable finding of this structure is the lack of direct information about the dissertation itself. For a document that concerns years of research progress by the student, it is interesting that

Requirement	#	Description
Research Credits	70	Credits named with connection to research connected to the dissertation
Naming	181	Formatting or presenting the dissertation
Timeline	61	Expectation of when the dissertation should be defended

Table 4.8: *The requirements of the Dissertation structure. Each structure includes the requirements and a brief description of the requirement determined during the coding process.*

Requirement	#	Description
Number of Faculty	99	Number of faculty required for committee
Types of Expertise	52	Specifics of expertise of the committee members. Includes whether members need to be outside student's research sub-field or physics department

Table 4.9: *The requirements of the Supervisory Committee structure. Each structure includes the requirements and a brief description of the requirement determined during the coding process.*

there is little about the writing and expected content of the document. For some programs, there is explicit language stating it is up to the committee to determine the formatting and content of the dissertation, or stating that the formatting must be in line with the regulations of the program's graduate college. Additionally, we note that the language for the dissertation of a reflection of the student's ability to do independent research has not changed significantly since the PhD was first introduced in 1861 with language that the dissertation is an 'exercise in scholarly research' [?] .

4.4.4 Supervisory Committee

The Supervisory Committee is composed of faculty who provide feedback to graduate students during their candidacy process and the dissertation defense. This group also is the ultimate decider in whether that student passes candidacy and the defense. This structure involves 2 requirements involving what faculty can be considered part of the committee. The number of programs that have each requirement is shown in Table 4.9

The Supervisory Committee is a graduate requirement that spans two separate parts of a graduate student's journey. It is named as its own structure, rather than Candidacy or Dissertation, as it spans both groups and, in many cases, subtly changes between the candidacy process and defending the dissertation. The supervisory committee largely concerns forming an advisory group to monitor a graduate student's progress during graduate school. The members of the committee include the student's chosen advisor and several other faculty members. 151 institutions listed the requirement for a Supervisory Committee, with 99 specifying the number of members required. The average number of committee members needed for Candidacy was just under 4 members, and the number of committee members needed after Candidacy for the Dissertation was just over 4. The expectations for these committee members often included when the group needed to be formed, how many members, and what the composition of the group needed to be. 52 institutions also require that one of the faculty members must be either outside the department, the research group, or the research sub-field of the student.

4.4.5 Training Modules and Courses

In total, 42 programs had explicit training requirements for their graduate students, with some of these programs having multiple training modules or courses. We identify 5 distinct requirements for the training structure: Teaching Assistants (TAs), Ethics and RCR, Safety, Human Subjects, and Hazardous Waste. The number of institutions listing each requirement is found in Table 4.10. We find them distinct based on the outcomes of the training and the overall goal of this type. First, concerns the training of TAs. This requirement included all opportunities for students to engage with training in teaching and communicating physics to students, and the naming of the training requirements often reflected this, including language like Pedagogical Training, Teaching Assistant Seminar, or Teaching of Physics. The range of time commitment for this group was diverse, with some acting as quick modules to be taken before semester started, to full courses taken for credit. The expectations were often that the students would have opportunities to learn about techniques for teaching and how to manage different situations. The number of institutions that explicitly mention the mode of the training requirement is listed in Table 4.11.

The second requirement concerns training graduate students in ethical practices. This group is largely concerned with training modules, sometimes outsourced to different groups, such as the

Requirement	#	Description
TA Training	27	Pedagogy training related to teaching
Laboratory Safety	3	Safety training related to their research work
Ethics/RCR	26	Students to receive research ethics training
Other	1	Other training opportunities required by the department

Table 4.10: *The requirements of the Training structure. Each structure includes the requirements and a brief description of the requirement determined during the coding process.*

Table 4.11: *The list of Training required from the total of 181 institutions. The table below includes the different types of training required, with the ones listed as modules only in column 2, the ones listed as full courses in column 3, ones not listed with a time frame in column 4, and the total number of institutions with that training type, T, in column 5.*

Requirement	Module	Course	Unlisted	T
Teaching Assistants	18	8	1	27
Ethics / RCR	17	5	0	22
Safety	1	0	1	2
Human Subjects	1	0	0	1
Hazardous Waste	1	0	0	1

CITI Program, to be taken online to satisfy the requirement. The subject of these training modules is largely concerned with research ethics education, and more specifically with understanding the 9 core principles of Responsible Conduct of Research (RCR). In some cases, institutions require students to take the online training modules from CITI as well as a full Pass/Fail course. In the event that students do not pass this course, they are then required to take a course within the Philosophy department at those institutions.

The third requirement is Safety. These training opportunities concerned training graduate students in proper laboratory safety, including expectations in the lab, working with equipment, and emergency protocol. The fourth is Human Subjects, which is primarily concerned with the safety and privacy concerns of research concerning Humans, and the fifth requirement concerns the handling and safety concerning hazardous waste. We identify these last three as distinct from the other groups as they potentially link with research interests.

Notably, the training modules described by this structure tend towards more general graduate education, such as the RCR training or TA training. We identified very few training requirements

Requirement	#	Description
Format	59	Written and/or presentation to committee
Destination	10	Department or college the annual evaluation is sent to

Table 4.12: *The requirements of the Annual Reports structure. Each structure includes the requirements and a brief description of the requirement determined during the coding process.*

that may be connected with a student's research activities. The more likely training requirements of this structure that could be applicable for a student's research development could be in Human Subjects, Hazardous Waste, or Safety. However, both of these requirements are required by only 1 to 2 programs in this data, so it is not clear how widespread training is for physics graduate students.

4.4.6 Teaching Assistant Requirement

This structure has only one requirement, whether a program requires all graduate students to be TAs. Despite being largely accepted that many graduate students would be TAs, only 45 institutions formally require it as part of their graduate programs. In many cases, students may be expected to be enrolled as Graduate Teaching Assistants (GTAs), but there are often opportunities for students to begin research right away. We would again like to highlight that a limitation of this study is the likelihood that the presented data may not match what happens at those institutions. We additionally highlight that students who are interested in attending graduate programs likely could not find any information about being a TA at that program.

4.4.7 Annual Reports

The Annual Report is a collective opportunity for faculty in the supervisory committee and program to check student progress each year, with 59 institutions specifically mentioning the practice. There are 2 requirements in this structure, and the number of programs listing each requirement is shown in Table 4.12. These requirements include the destination of the report itself, including the department or the graduate college. The students will continue to fill out these reports each year until the defense of their dissertation.

Requirement	#	Description
Venue	21	Location of presented posters or talks either in the department or outside
Format	28	Format of presentation including posters or talks

Table 4.13: *The requirements of the Presentations of Research structure. Each structure includes the requirements and a brief description of the requirement determined during the coding process.*

The way that programs discuss the annual report varies. Students may present their progress that year in a formal talk to the full committee, get feedback on that progress, and the report is signed and sent to the department or graduate college. Often, this structure is described as the student making a report to their committee each year, with no specifics on the formatting or presentation of that report. Something of further analysis could be to investigate how programs consider the value of the annual report and what function it serves in the development of research understanding for the student.

4.4.8 Presentations of Research

The last structure concerns any presentation of research made by the student. This structure has 2 requirements that involve both the format and venue of the presentation. In the course of completing a graduate degree, students engage with research and may attend conferences, colloquia, or workshops to present talks or posters. These kinds of presentations are fairly standard in the norms of graduate education, yet physics graduate programs rarely require students to formally present progress in research. The most common requirement of this kind is to have students present some kind of work at a departmental colloquium, at 17 programs. We note here that the requirements of students submitting papers for conference proceedings or presenting posters of any kind are very low, at 2 programs. The last requirement of this kind we found was for students to be published in an academic journal before defending. With the amount of writing, potentially, required for the dissertation, we find only 7 programs requiring that students publish.

4.5 Discussion

This study is aimed at addressing two aspects of physics graduate education. First, to determine the variability of program structures in physics graduate education in the United States, and then to determine how these structures inform varieties of approaches in graduate physics curriculum. The structures of physics graduate education have an impact on both the graduate students and faculty. The ways these structures are coordinated as part of the graduate physics curriculum allows us to understand the different approaches that programs across the country. In the framework of structure-deficit, we identify that the present structures of physics graduate education as potential barriers for students. This work will first discuss the holistic curriculum of physics graduate education built from these structures and then the way that particular structures interact with each other.

4.5.1 The narrative of physics graduate education

The NASEM report identifies a process of graduate education as:

> ...while no two PhD experiences are identical, the PhD programs typical of many STEM disciplines includes 1 to 2 years of discipline-specific coursework; perhaps one or more years as a TA; the search for a dissertation advisory; which may or may not involve a formal system of rotating through several laboratories; comprehensive subject matter examinations; formulation and defense of a dissertation project; 3 to 7 years of mentored research supported by a combination of research assistanceship and fellowships; writing the dissertation; and a final defense of the dissertation[4].

The structures of physics graduate education described in Section 4.4 were determined based on the data from 181 physics programs of PhD-granting institutions. We determined these structures based on the prevalence of each structure's requirement in the publicly available data of each program. Considering all structures together, many of them have requirements involving a particular time frame in which the structure operates.

The Core Courses, Elective Courses, Comprehensive Content Exams, Training, and TA Requirements structures have both explicit and implied characteristic requirements. For example, the 9 programs that have a timeline requirement for the Core Courses structure often require that students finish their core courses within the first two years of their program. If we consider the first two to three years of physics graduate education, we identify a strong emphasis on students gaining physics content knowledge. Students are expected to take some particular courses or subjects in physics. Programs additionally impose characteristic requirements that concern the courses, including the timeline and grades required. Students move from general physics content knowledge to more specialized content knowledge through elective courses. At the end of 1 to 2 years of graduate education, many physics students take part in candidacy exams. A common structure of candidacy exams is the Comprehensive Content Exam. The written format of the Comprehensive Content Exam often focuses on physics content knowledge. The level described is often 'upper-division to lower graduate levels' of physics content knowledge.

For many institutions, Candidacy also involves the student being assessed on their current understanding of research practices, such as writing and presentation of any progress they have made in a research problem. The process of students becoming PhD Candidates often takes place 2-3 years into their programs. The time-sensitive requirements of structures after Candidacy are more spread out. The annual report may include a yearly presentation of research to the supervisory committee. The student may take research credits associated with the dissertation each year. The last time-sensitive requirement is the maximum number of years a student can stay in their program, coinciding with the defense of the dissertation. The requirements of the structures present for PhD candidates tend towards research preparation for the dissertation. Part of the candidacy process for many programs involves a presentation involving the student's research, which continues to build each year, assessed each year by the supervisory committee, students engage in research credits, until the student defends. The years following candidacy are also less universally organized. Before candidacy, every student in a program likely takes the same courses, but after candidacy, students move from the cohort arrangement of students towards engaging more fully with their own research groups. To accommodate all possible research groups' progress and norms, the structure requirements reflect this accommodation.

The common structures identified in this study are representative of what the community expects from graduate education. Some noted concerns from the literature of graduate education include the lack of preparation towards writing the dissertation[89] and feeling not properly prepared for the job market beyond graduate school[28–30]. Writing Courses are an uncommon requirement in the common structure of Core Courses. Scientific writing was listed as a Core Course in 2 programs in the data. Scientific writing is a critical practice expected by graduate students in several structures, namely the written format of the Research Readiness Exam and the Dissertation itself. There are several practices, like writing or choosing a suitable research group, that are not common requirements in this dataset. In the apprenticeship model, the student, as the novice, gains specialized knowledge and experience from the experts in the small learning community. If we consider this model, the student would learn the particular norms of the research group they join, and practices such as writing norms may be learned here. Understanding where students learn these particular practices is an avenue of further analysis.

4.5.2 Considering how physics graduate programs prepare and assess student progress

A critical element of physics graduate education is the training of graduate students towards becoming experts in their particular subfield. In this study we identify a divide in the way that physics graduate programs consider between content knowledge and research preparedness.

The preparation and assessment of physics content knowledge

The beginning of many graduate programs emphasizes the preparation and assessment of physics content knowledge. The Core and Elective Courses structures are well-defined and common in this study. The vast majority of core courses are identified by name, and the requirements of the elective courses are defined by the number of courses or credits required. If the courses are expected to be taken by a particular time, the institution will specify that it done by the first few years. In the handbooks and websites, the courses were often the first piece of information provided.

Often immediately following the information about courses in the data was the process of candidacy. A major structure of candidacy is the Comprehensive Content Exam. The purpose of this candidacy exam is to assess a student's knowledge of physics content, often to the level of 'upper-division to lower graduate' understanding. It most often takes the form of several exams on different topics, often covering the subjects of Classical Mechanics, Quantum Mechanics, Electrodynamics, and Statistical Mechanics. For those institutions where the subjects of the Core Courses align with the subjects of the Comprehensive Content Exam, this provides an additional way for programs to assess if students have received enough preparation to reach the level of physics understanding deemed necessary by the program. In any case where the program may require both Core Courses and the Comprehensive Content Exam, further analysis will be needed to determine if the Comprehensive Content Exam is redundant with the final exams of the required Core Courses.

The most common subjects of the Core Courses and Comprehensive Content Exam structures fall into a classification of, say, 'general' physics. Where the program determines that every graduate student in the program must be at to earn a PhD, accounting for different sub-disciplines in physics research. The role of the Elective Courses, as sometimes described in the data, is to expand the students' knowledge in breadth beyond these more general courses. The requirements of the Elective Courses structure are also slightly different than the Core Course structure. These requirements focus more on the number of courses or credits required, or by providing the student a list of courses in particular sub-fields, like astrophysics or plasma physics, to choose from. These courses allow students to become more specialized in the subfield, potentially aligned with their research interests. It is currently unclear where this specialized knowledge may be assessed in isolation from any research progress. The verbal format of the Comprehensive Content Exam is not as clearly defined by the data, and may involve the committee supervising the verbal exam to ask questions related to the student's specialized knowledge. Another possible point where this knowledge could be assessed is through the Research Readiness Exam. Depending on the format chosen by the program, this may involve the student expressing any specialized knowledge during the Candidacy Exam. Ultimately, any decision to have the student do so will be entirely up to the Supervisory Committee.

The assessment of research understanding

Physics graduate programs have a defined way to assess the progress students make in research understanding but limited ways for programs to prepare students to gain this understanding. The Research Readiness Exam is fundamentally an assessment of the progress a student has made in research, either through a proposal towards the dissertation or a literature review of a topic related to their interest. The ultimate assessment of the progress a student has made in research understanding is in the Dissertation itself, reflected in the way the document itself is considered as a 'piece of scholarly work completed by a student'[26]. The Annual Report is a consistent way to assess the progress a student is making between these two major milestones (Candidacy and the Dissertation). The Annual Report is an uncommon structure in this data, so the breadth of ways the Annual Report is implemented is not clear. As discussed above, the role of the student in graduate education shifts from novice to expert researcher by the defense of the dissertation. The structures of the Research Readiness Exam, Dissertation, Annual Reports, and the Supervisory Committee all help to monitor the student's progress toward the Dissertation Defense and provide feedback at every level.

The assessment of research knowledge and progress is well established in physics graduate education; the preparation and training for it is not. Within the common structure of Core Courses, any requirements that may impact research progress, such as professional development, Research Rotation, or Scientific Writing, are rare, with less than 10 courses identified. Additionally, the uncommon structure of Training has limited requirements that could impact training. In particular, the training in Responsible Conduct of Research takes multiple forms in this study. The overall effectiveness of most forms of RCR training is questionable, and any lasting impact this training may have on the training of Graduate Students is unclear. There are limited opportunities for graduate students to be formally trained or prepared for the intensive research required for the dissertation.

4.5.3 The different assessment goals of candidacy exams

The two structures of Candidacy Exams appear to accomplish different goals, each providing barriers for student progress. The Comprehensive Content Exam is an assessment of a student's understanding of physics to a very particular level determined by the community and program. The Research Readiness Exan is an assessment for programs to understand where the student is in their understanding of research practices, and often the first explicit point of feedback for students from their Supervisory Committee. There are two formats for each of the Candidacy Exam structures: a written and/or a verbal type. We identified 10 distinct ways that programs in this data approach the Candidacy process. An avenue of further investigation is to understand the value each way brings to its program. The Comprehensive Content Exam is often the final assessment of more 'general' physics content knowledge, and the Research Readiness Exam is an opportunity for students to receive feedback towards advancing their research towards the Dissertation.

Candidacy Exams are a noted point of concern in the literature. The student perception of Candidacy Exams is often poor, often viewing the practice as little more than a trial by combat[?] or as a means of 'weeding out' students[97;98]. The two structures of Candidacy in this study each offer some troubling implications. The common subjects covered in the Comprehensive Content Exam are also the common subjects of the Core Courses structure. This overlap may prove to be redundant for some programs. Students are often required to maintain a certain G.P.A. in their core courses, and to do so, must pass these classes through some course assessments. Those same students may then be required to take another exam covering some form of those same subjects to a satisfactory level. The Research Readiness Exam is largely an assessment of whether a student is prepared for the work necessary for a dissertation. We identified some significant variety in the way this structure is discussed in graduate programs, and often how little information may be given to students about the preparation for this assessment, which may be encouraging this perception of faculty apathy in preparing students[86]. Further analysis will be needed to make sense of how the different paths of candidacy have been chosen by programs, how these candidacy exams are meeting the goals of the program, and the impact that each path has on student progress through their graduate programs.

4.5.4 Aligning recommendations with structures of graduate education

Concerns of graduate education are well-documented, and the reports from NASEM propose recommendations that address those concerns are enumerated above. The National Academies proposes recommendations towards building an 'Ideal Graduate STEM Education'. The following will be a cursory step towards understanding if the structures of physics graduate education can be used to assess if programs are meeting the recommendations of this report. The two recommendations considered in this section are part of a list of 56 total recommendations for various levels of graduate education. First,

> Graduate departments should publicly post how their programs reflect the core competencies for doctoral students, including the milestones and metrics the department and individual faculty use in evaluation and assessment[4].

The entirety of this study is based on public data that is posted by each department. Through this data, we determined 10 major structures of physics graduate education with numerous requirements. Clearly defined structures include Core Courses, Elective Courses, Comprehensive Content Exams, and Training. Each of these structures has numerous requirements that are generally well-defined, such as the courses required or when the Comprehensive Content Exam must be taken. Other structures are less clearly defined, such as the Research Readiness Exam or the Dissertation, as discussed above. This public data is necessary for various groups to make sense of the requirements and expectations of each practice from the department, which may be critical for both students and faculty. A lack of critical information or guidance could make it difficult for students and faculty to progress through particularly unclear practices.

Another recommendation from NASEM reads,

> Institutions should integrate professional development opportunities, including relevant course offerings and internships, into doctoral curriculum design[4]

In the structures of this study, we identified few requirements that could relate to the professional development of students. In the Core Courses structure, we identified very few courses that align with this recommendation. We did identify core courses such as Professional Development or Best

Practices in College Teaching. These courses were not common in this structure, required by 5 programs or fewer per course. The socialization of graduate students towards becoming an expert in their research interest is a pillar of graduate education. Students likely receive their professional development opportunities from their research groups and advisors. This model does not currently allow programs to continue supporting the overall professional development of students, particularly in the first few years of graduate school. Another element of this concerns the present perception of the 'job crisis', where students are not feeling prepared for the jobs they take. In the present structures, there are no explicit requirements that discuss the professional development of students in terms of career exploration or preparation.

These two recommendations are only a small portion of the extensive list produced by the report from NASEM[4]. By introducing these two recommendations, we find that the structures can be discussed in this context; however, a more critical and thorough analysis of the way the present structures align with these recommendations is needed to determine the strength of the present graduate education system.

Recommendations to departments

In light of the structures present in this study being constructed from the public program requirements, we would like to propose our own recommendations.

First, physics graduate programs should articulate how requirements align with their learning goals to both students and faculty. There are programmatic requirements of physics graduate education that are well discussed, such as the courses required and the training, when required. Some elements, such as the process of candidacy and the dissertation, are often not described in detail, leading to well-established concerns from both researchers and graduate students. An analysis of identifying the learning goals of graduate education has already been conducted in Chemistry, finding that clearly stated learning goals allowed for departments to better align assessment tools to those goals[88]. Doing so may help to reduce confusion among students and faculty, as seen in the faculty perceptions of candidacy exams in physics[47]. Clearly stating these goals and potentially aligning structures with those goals could articulate to students and faculty the expectations of the department for those structures and how they will be assessed.

Departments should ensure that there are structures in place to prepare students for significant milestones. This model already exists for several structures in this study. Students gain significant content knowledge and are assessed on that knowledge, potentially multiple times depending on the structures of the program. Other structures may have limited requirements that prepare for those assessments. We see this in the Research Readiness Exam, where this candidacy exam is an assessment of research knowledge and understanding. There were a few requirements concerning the research development of students, as they were often in uncommon core courses. Students between candidacy and the dissertation are expected to become more expert-like in research understanding, but to the department, this can only be assessed by the annual reports and the dissertation itself. Departments should consider what structures could be put in place to support the development of research understanding for graduate students throughout their journey.

Departments should evaluate their doctoral curriculum to identify redundancies and where structures may not be supporting one another. Graduate students in the first few years of a program engage heavily with learning and being assessed in content knowledge. This content knowledge may be assessed multiple times during the graduate student's time in the program. Departments should consider whether the assessments from the Core Courses structure and the Comprehensive Content Exam structure are redundant. Additionally, there may be a disconnect between the process of Candidacy and the defense of the Dissertation. In particular, programs should consider if the assessments of the Candidacy Exams are preparing students for the work necessary for the dissertation, or if some stronger connection could be made in terms of skills and knowledge gained during candidacy that could be applied to the dissertation defense.

Chapter 5

Physics graduate student retention and degree completion rates

5.1 Motivation and outline of this chapter

Student retention is a common measure of the success of a program. Institutions greatly value this measure, and national organizations such as AIP regularly publish retention statistics. At the graduate level, the importance of retention is similarly critical. Retention at the institutional level allows a program to begin understanding how the program is affecting its students, and promote further investigations into the students' experiences, ultimately towards increasing this rate. Previous studies of retention/attrition/persistence of graduate students have either focused on particular students, or cohorts of students, or national information on graduate students. The mode of prior methods has focused on modeling student retention while centering the experiences of the students to determine reasons students may or may not persist in their programs. This study then attempts to determine the feasibility of constructing measures of retention and completion rates of physics graduate students for individual programs through enrollment and degrees awarded data. The research questions that will be guiding this chapter are:

- Can a measure of yearly graduate student retention be built from student enrollment and degrees awarded data?

- Can a measure of degree completion rates be built from student enrollment and degrees awarded data?
- Are these measures feasible for individual physics graduate programs?

Section 5.2 will consolidate the 20 years of AIP survey information towards the data used in this study. The construction of each measure is described in Section 5.5, and the application of these measures to national and institutional data is found in Section 5.6. The feasibility of these measures is described in Section 5.7.

5.2 AIP survey data

The data for the entirety of this study comes from the yearly reports from the American Institute of Physics (AIP)[123]. These reports represent the enrollment and degrees granted by each institution in the United States. The information is self-reported from each institution, and the report states both the total enrollment and degrees granted across the nation and from each school individually. The data collected includes institutions ranging from exclusively granting bachelor's degrees to graduate degrees and certifications in physics, applied physics, medical physics, etc. This data set provides a near population-level measure of graduate student enrollment and degrees granted in the United States over the past few decades at an individual institutional level, providing a unique longitudinal opportunity to study graduate student retention on a macro and micro scale. While AIP regularly publishes analyses of this data to describe trends in enrollment and degrees awarded[123] in physics and astronomy, to our knowledge, such data is not used to measure retention or completion rates in physics graduate education.

These reports have been collected and aggregated for decades, but for this study, we focused on a 20-year window from 2000 to 2022. We collected these reports in the Spring of 2024 and adjusted the growing data set to account for institutions adding or dropping the different degrees awarded in physics, and those institutions that changed names between 2000 and 2022. The reports each average around 750 individual institutions. We constrain this data set to cover institutions granting graduate degree types. This would then exclude all bachelor's-only institutions and those that offer Master's certificates or accelerated Master's degrees, as those institutions may not have their graduate programs. Doing so reduced the full data set to 284 institutions. This full data set includes 214 PhD-granting institutions and 70 Master's-granting institutions.

5.3 Defining variables

This work intends to determine a retention measure and degree completion rates for physics graduate students. The way that we define Retention, then, is the percentage of students present in their department each year. Several factors impact this definition, including the number of students being awarded their degrees and those entering the program. The Completion Rate of a population of students includes the number of degrees awarded to a particular cohort, or set of cohorts, of graduate students. This implies a time scale, where the number of first-year students is compared to the number of degrees awarded some time later. The data from AIP is reported in the fall term of each year, and this provides us with a baseline to compare the data from each survey year to year. The data present in the study includes both the relevant information of each data point and the time period of the collection.

The data that we are interested in includes the variables described in Table 5.1. In particular, we are interested in 4 categories of data from the surveys. These categories include: Total Graduate Students, First Year Graduate Students, Exiting Master's Physics Degrees, and PhD Physics Degrees. The time scales associated with these categories are also described in these roster surveys. For example, in the data from 2022, the First Year variable (F_{2022}) would consider the first year graduate students who were new to the department in the Fall of 2022, as well as the students who entered in the previous winter, spring, and summer terms. The only variable with no associated time frame is the Total Student variable. We propose that the collection period for the Total Graduate Student variable takes place each fall, as it aligns with the collection period for the roster survey itself, according to AIP and the solicitation for all programs. In addition to the data presented in the survey, we propose an additional variable. We must account for the number of students who are not enrolled between the Fall terms of two years. We name this variable N_n, where the number of students, N, not enrolled in calendar year n.

Some notable limitations of these variable definitions involve the Master's degree. The variable

Table 5.1: *Categories of student enrollment data from AIP survey data collected in Fall of each year. Each variable described includes the specific demographic of students included and the time period of collection used.*

Variable	Name	Description
F_n	First Year Graduate Student Enrollment	Number of students who started Graduate School in *calendar* year n.
P_n	PhDs Awarded	Number of PhDs awarded during the *academic* year beginning in Fall of year n.
M_n	Exiting Master's Awarded	Number of terminal Master's degrees awarded during the *academic* year beginning Fall of year n.
T_n	Total Graduate Students	Number of Graduate students enrolled in Fall semester of year n, including first year and degree recipients.
N_n	Non-Enrolled Students	Number of students not enrolled by fall semester of year n

M_n only considers the exiting Master's degrees awarded. Any Master's that a graduate student is awarded en route to the PhD is not captured in any variables described in this study.

5.4 Time scales and counting students

The variables described in Table 5.1 include both the types of students included in each variable and the time for which the information was collected. Something that must be done before building any measure of student retention or completion rate is to ensure the variables for this study correctly account for changes in student populations each year, specifically for students who enter and leave their program at different points in the calendar year. To check this, Figure 5.1 is a student trajectory map that includes the Spring, Summer, and Fall terms of two calendar years $(n-1)$ and (n). This includes students leaving the program each semester with degrees and not enrolling the next semester, those continuing in their program between the two years, and the first year students enrolled in each term. We place the variables from Table 5.1 in their correct collection periods on the right side of this figure. An alternative view of the different time scales used in Table 5.1 can be seen in Figure 5.2.

Considering both Figures 5.1 and 5.2, we propose that the change in the population of graduate

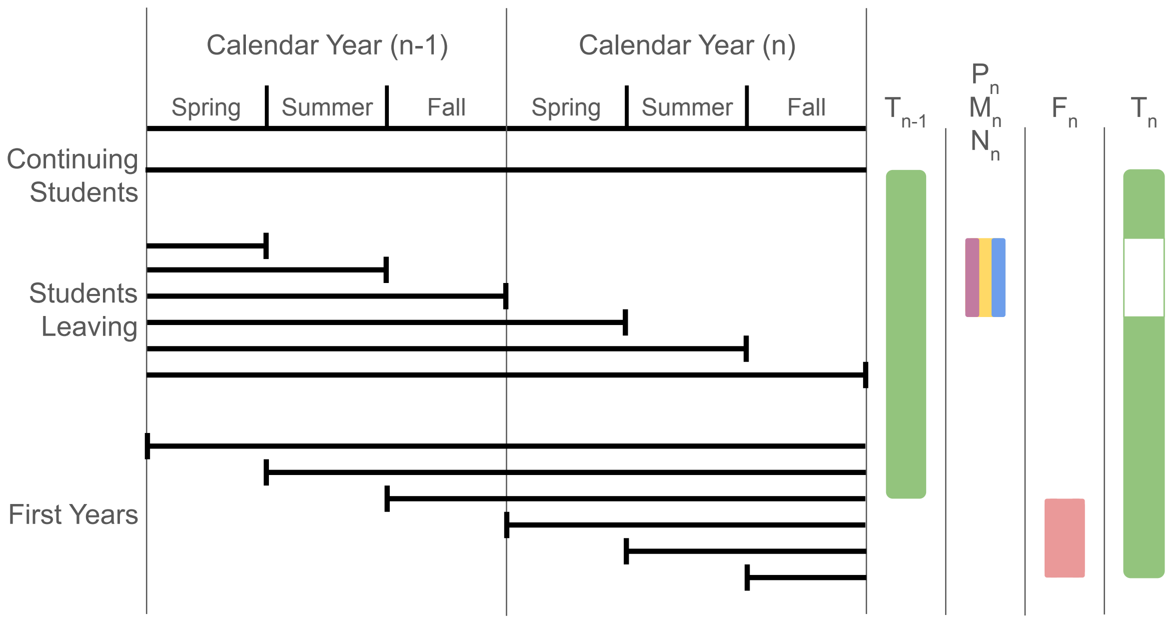

Figure 5.1: *A mapping of graduate student trajectories in their program across the Spring, Summer, and Fall terms of two calendar years* $(n-1)$ *and* (n)*. The left includes the points when students would be leaving their program through degrees being awarded and not being enrolled between terms. It also includes the students continuing their program across these two years, and those students enrolling in the program for the first time in each term.*

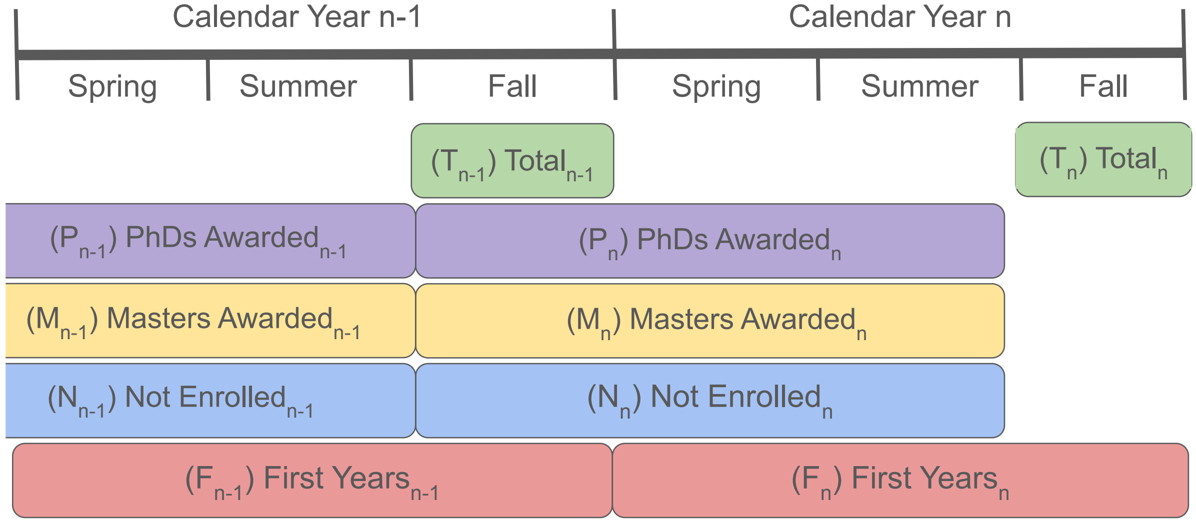

Figure 5.2: *A Gantt Chart of the Fall, Summer, and Spring terms of two calendar years* $(n-1)$ *and* (n)*. The variables used in their study are placed in their respective place in the Chart.*

students is determined by considering the students persisting in their program, balanced with the students entering and leaving. The students who persist in the program are joined by the first year students who enter are different points in the year between Fall Terms. We define this balance in Equation 5.1 utilizing the variables from Table 5.1.

$$T_n = T_{n-1} - P_n - M_n - N_n + F_n \tag{5.1}$$

5.4.1 Adjusting data for analysis

The analysis for this study is interested in developing measures of retention and completion rates for both national and institutional data. The 284 graduate institutions of the dataset contain graduate institutions that are both PhD-granting and Master's-granting programs. Considering the 284 institutions between 2000 and 2022, there are 5667 points of data that are not filled in, equal to 22% of the total possible points of data. Master's-granting institutions would not have any recorded PhDs awarded for any years. The remaining points are not filled for reasons currently unknown to the authors.

We impose two different scales of data for this study. First, on the National Level, consider the totals of each variable for every year. For example, if we wish to consider the number of PhDs awarded in 2006, then we must total the PhDs awarded from all universities that reported this data in 2006. Considering only the national data between 2000 and 2022 reduces the data to 92 points of data, with 4 variables for each year. Next is the Institutional Level, which considers institutions with complete data points. Doing so further constrains the institutions available for analysis to 60 PhD-granting institutions and 7 Master's-granting institutions.

5.5 Developing measures

The following section will describe two sets of measures. First, we will construct a yearly retention rate for physics graduate students. This retention rate is based on the changes in student populations described in Equation 5.1. The primary yearly retention rate of Equation 5.2 describes the changes, including both Master's and PhD Students. Equation 5.4 describes a variation of this rate that considers the retention of only PhD Students.

The second set of measures for this study will focus on completion rates for degrees awarded to cohorts of first year graduate students. Equation 5.5 describes the completion rate of graduate students awarded a PhD, assuming a time scale of six years after first enrolling. Equation 5.6 describes the completion rate of students awarded a Master's degree, assuming a time scale of two years, and Equation 5.7 considers a completion rate of students awarded either a PhD or a Master's based on the same cohort of first year graduate students.

5.5.1 Graduate student retention rates

The first measure we introduce focuses on the yearly retention of graduate students, comparing the number of students entering and leaving their graduate program. As we are focused on the retention of all physics graduate students, compared to the total population, we then define the Graduate Retention Rate, GRR_n, of a particular year n as:

$$GRR_n = 1 - \frac{N_n}{T_{n-1}} = \frac{T_n + M_n + P_n - F_n}{T_{n-1}} \tag{5.2}$$

Utilizing Equation 5.1, we define N_n in terms of the variables from Table 5.1. We propose that the percentage of students not enrolled based on the total population of the previous year is $\frac{N_n}{T_{n-1}}$. In this same frame, there must additionally be a percentage of students retained in the program, here defined as GRR_n. Critical to this analysis is the presumption that the percentage of students non-enrolled ($\frac{N_n}{T_{n-1}}$) and the percentage of students retained (GRR_n) in that same time frame should be equal to 100%.

PhD graduate student retention

An alternative version of Equation 5.2 considers the retention of only PhD students. We start this from a similar position as above, as we consider the number of students present in the graduate program each year. To directly measure the changes in PhD student populations, we propose a variation in the variable for the students not enrolled as $(MN)_n$. Where the number of students not enrolled each year now includes the students who may be awarded a Master's degree. Considering this new value for students not enrolled, we construct the following equation for the changes in Total Student Enrollment (T_n), now only considering the number of incoming First Year students (F_n) with the number of PhDs awarded (P_n).

$$T_{n-1} - P_n - (MN)_n + F_n = T_n \tag{5.3}$$

Following the construction as above, we describe the following equation in similar terms. The notable difference here is the absence of the Master's awarded. This measure carries a critical assumption, that any Master's degrees awarded are inherently part of the number of students not enrolled each year. The First Year student population, as reported by AIP, does not determine a difference between students declared as PhD students and those declared as Master's students. Equation 5.4 describes the PhD Student Retention Rate ($(PRR)_n$) retained in their graduate programs.

$$PRR_n = 1 - \frac{(MN)_n}{T_{n-1}} = \frac{T_n + P_n - F_n}{T_{n-1}} \tag{5.4}$$

5.5.2 Graduate degree completion rates

PhD completion rate

Next, we propose a way to measure the rate of students who earn a doctoral degree in physics. The two variables that we wish to focus on are the PhDs being awarded (P_n) and the First Year student enrollment (F_n). This allows us to directly compare the degrees awarded with the students who enter their graduate program. A result of the AIP/AAPT task force report stated that the average time it takes for graduate students to earn a PhD in physics is 5 to 6 years[33]. To fully capture the number of first year graduate students who earn a PhD, we will need to compare the number of PhDs awarded to the appropriate cohort year of first year student enrollment.

Considering the expected time that graduate students earn a PhD, we propose the following Equation 5.7. Here, PCR_n (PhD Completion Rate) is the number of PhDs awarded (P_{n+6}) six years after the cohort of First Year students (F_n) first enrolled. While there may be an average time scale for students to be awarded a PhD, it is not inclusive of those students earning degrees outside this time frame. To account for this, we impose a scaling factor (x) as a way to involve more data for each year's calculation. The two values of x chosen for the rest of this paper will be $x = 0$ and $x = 1$. In the case when $x = 0$, then the expression returns a direct comparison between the PhDs awarded six years $(n + 6)$ after the First Year enrollment of year n. For example, calculating the PhD Completion Rate of 2012, or $PCR^{(0)}_{2012}$, would compare the PhDs awarded in 2018 with the First Year cohort of 2012. In the case of $x = 1$, the expression then averages each variable with the surrounding years. For example, if calculating the PhD Completion Rate in 2010 or, $PCR^{(1)}_{2010}$, then the expression would average the PhDs awarded in 2015, 2016, and 2017; then compare it to the average First Year enrollments of 2009, 2010, and 2011. We propose that the variation of this measure to average degrees awarded and student enrollment would further emphasize a way to calculate the completion rate of graduate students without having to track particular cohorts of students.

$$PCR_n^{(x)} = \frac{\sum_{i=-x}^{x} P_{n+6+i}}{\sum_{j=-x}^{x} F_{n+j}} \tag{5.5}$$

Master's completion rate

Next, we propose a way to measure the rate of students who earn a Master's degree in physics. The two variables that we wish to focus on now are the Master's being awarded (M_n) and the cohort of First Year student enrollment (F_n). From another analysis completed by this group, Master's degrees are often awarded within the first two years of the student's first enrollment in graduate school[124]. To fully capture the number of first year graduate students who earn an exiting Master's, we will need to similarly compare the number of Master's awarded to the appropriate cohort of first year student enrollment. This is most useful for Master's-granting institutions, where the Master's degree is the only graduate degree offered. For a PhD-granting institution, we identify no way to distinguish the First Year enrollment between students seeking a PhD or a Master's.

Considering the expected time that graduate students earn a Master's, we propose the following Equation 5.6. Here, MCR_n (Master's Completion Rate) is the number of Master's awarded (M_{n+2}) two years after the cohort of First Year students (F_n) enrolled. We impose the same scaling factor, (x). The two values of x will similarly be $x = 0$ and $x = 1$. In the case when $x = 0$, then the expression above returns a direct comparison between the Master's awarded in year $n + 2$ with the First Year population of year n. In the case of $x = 1$, the expression then averages each variable with the surrounding years. For example, if calculating the Master's Completion Rate in 2006 or, $MCR_{2006}^{(1)}$, then the expression would average the Master's awarded in 2007, 2008, and 2009; then compare it to the average First Year enrollments of 2005, 2006, and 2007. We propose that the variation of this measure to average degrees awarded and student enrollment would further emphasize a way to calculate the completion rate of graduate students without having to track particular cohorts of students.

$$MCR_n^{(x)} = \frac{\sum\limits_{i=-x}^{x} M_{n+2+i}}{\sum\limits_{j=-x}^{x} F_{n+j}} \tag{5.6}$$

Degree completion rate

By investigating the degree completion rates of cohorts of first year graduate student enrollments, we can additionally combine the previous equations to capture the completion rate of both degree types. As both Equations 5.5 and 5.6 have different time scales associated with them, by centering each equation on the first year population, we can construct an equation of both degree completion types. Equation 5.7 then describes the Degree Completion Rate, or $DCR_n^{(x)}$, which operates on the same scaling factor. As first year students seeking a Master's or a PhD is not specified, by centering this measure on particular cohorts of first year students, we suppose this measure would capture the percentage of First Year students who earn a graduate degree in general, without distinguishing which group of First Year students earn either degree. This measure is best applied to a PhD-granting institution.

$$DCR_n^{(x)} = PCR_n^{(x)} + MCR_n^{(x)} = \frac{\sum\limits_{i=-x}^{x} P_{n+6+i} + \sum\limits_{j=-x}^{x} M_{n+2+j}}{\sum\limits_{k=-x}^{x} F_{n+k}} \tag{5.7}$$

5.6 Applying the measures to the AIP data

The following results will be presented in two parts: considering the yearly retention of graduate students using Equations 5.2 and 5.4, and then the completion rates of graduate degrees using Equations 5.5 - 5.7. Data from 284 institutions were used in this study, with 4 variables of data used from the AIP surveys: Total Graduate Population (T_n), the First Year Enrollment (F_n), the PhDs Awarded (P_n), and the Master's Awarded (M_n).

5.6.1 A yearly graduate retention rate

The first measure was constructed to determine a yearly measure of graduate student retention. We consider first the measure applied at the National Scale, where the measure considers the totals of each variable between the years 2000 and 2022. In an effort to understand the measure as a whole, we will consider all resulting data points together rather than individual points. If we consider the results of this first measure applied to the National Scale, we find a range of retention rates between 86 and 117%. Figure 5.3 shows the results of this measure applied to the National Scale, totaling the data from 284 graduate institutions.

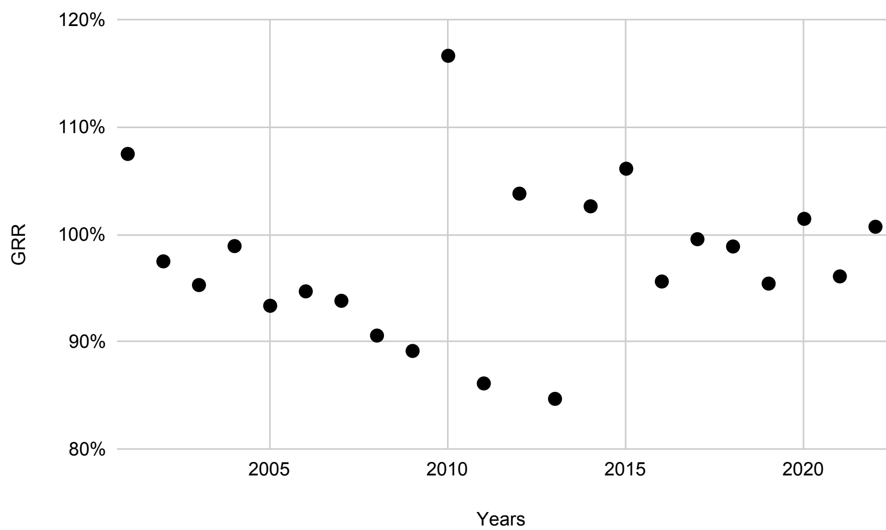

Figure 5.3: *The Graduate Retention Rate GRR_n applied to the National Level that totals the data from 284 graduate institutions. The value at each point considers the GRR_n of each year n starting in 2001.*

This measure can also be applied to determine the yearly graduate retention at Master's-Granting Institutions. This measure, applied to the National Level, totals the data from 70 Master's-Granting institutions and is shown in Figure 5.4. By inspection, it is difficult to make direct comparisons between these Figures 5.3 and 5.4. Overall, the yearly retention of graduate students appears to fluctuate, and it is difficult to determine any significant trends in each result at this time.

Now applying this measure on the Institutional Scale, we consider two PhD-granting institutions with complete data between 2000 and 2022. The application of this measure is found in

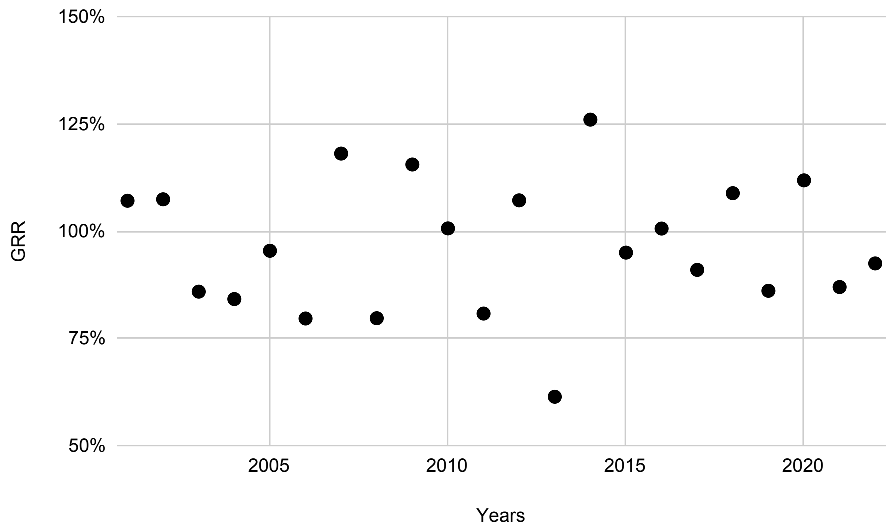

Figure 5.4: *The Graduate Retention Rate GRR_n applied to a National total of 70 Master's-Granting Institutions in year n. This considers the Graduate Retention as predominantly focused on the retention of only Master's Students.*

Figure 5.5. The two institutions were chosen at random, yet highlight that the retention rate of graduate students is different for each institution, as expected. This is observed for each institution where the measure was applied, with no particular sets of institutions or years that show particular trends in retention rates.

A notable feature that is present in Figures 5.3, 5.4, and 5.5 is a retention rate over 100%. Section 5.7 will discuss potential reasons for this phenomenon and the impact it has on this study as a whole. A yearly retention rate is something we would consider to not be realistic, as it would be driven by students 'showing up' each year that were not being counted before. If the rates are over 100% for multiple years, as it does in each figure, this suggests at this time that something is flawed in the construction of this measure, so we will not make claims on the findings of this measure yet before further analysis.

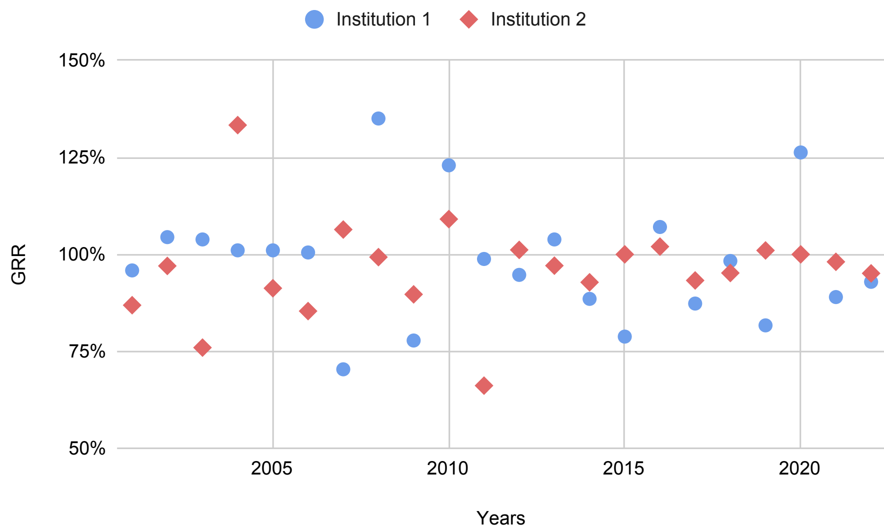

Figure 5.5: *The Graduate Retention Rate GRR_n applied to two PhD-granting institutions' dataset.*

5.6.2 Completion rates of first year physics graduate students

Degree completion rate

The second set of measures concerns the completion rate of first year graduate students. We first consider this measure as applied to the National Scale, totaling the data from 284 institutions. Each of the completion rate equations is built by comparing the number of degrees awarded to cohorts of first year graduate students. The scaling factor was introduced to attempt to capture a more complete representation of when graduate students are awarded their degrees. The Degree Completion Rate (DCR) is a combined measure including both PhDs and Master's degrees awarded, with each degree awarded in the time frame determined by the literature. Figure 5.6 shows the DCR applied to the National Level data set at both scaling factors, highlighting that the scale factor ($x = 1$) is smoother by inspection, as expected from taking averages of the degrees awarded and first year enrollment data. For the purposes of the rest of this paper, we will be reporting data only using the scale factor ($x = 1$). In both cases, we find that the percentage of graduate students being awarded degrees appears to be increasing.

Considering the $DCR^{(x=1)}$ for the Institutional Level, we consider the two PhD-Granting institutions with complete data points between 2001 and 2021. The results of the measure applied

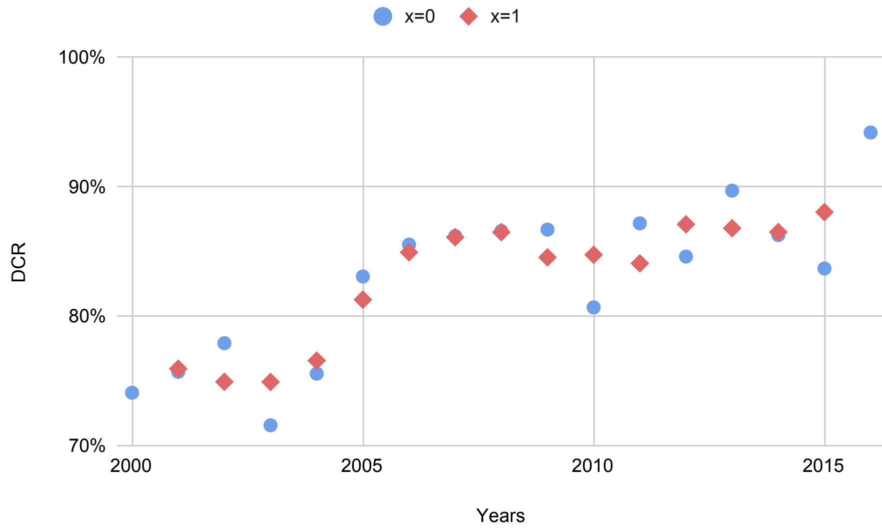

Figure 5.6: *The Degree Completion Rate DCR_n applied at the scaling factors of $(x = 0)$ and $(x = 1)$ applied to the National Data of all 284 graduate institutions in each year n starting in 2000.*

to these institutions are found in Figure 5.7. We see that each institution has a set of results, with no immediate connection to be made in trends between the institutions. The results of these institutions are representative of the application of this measure to all 60 full data PhD-granting institutions. We note again that the purpose of these measures at the Institutional Level is to consider trends in the results and not isolated points.

PhD completion rate

For the completion rate of PhD students, we consider the National Level of data from 214 PhD-granting institutions. This was done to minimize any potential effect the first year enrollment of Master's-granting institutions might have on this measure. This measure considers the percentage of first year graduate students who are awarded a PhD between 4 and 8 years later, as described by Equation 5.5. Figure 5.8 shows the measure applied to the National Level at a scaling factor of $(x = 1)$. By inspection, the rate of first year students who are awarded a PhD appears to be increasing between 2001 and 2015. Looking at the same two PhD-granting institutions with complete data points, Figure 5.9 shows the results of this measure between 2001 and 2015 as well. We find results distinct for each institution, suggesting that the measure is able to resolve at both the

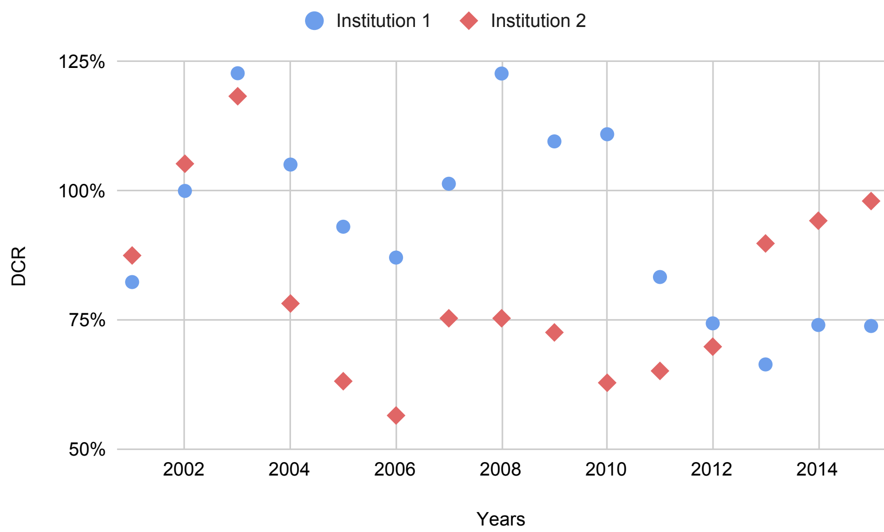

Figure 5.7: *The Degree Completion Rates DCR_n applied at the scaling factor of $(x = 1)$ to two PhD-granting institutions will have complete data spaces in each year n starting in 2001.*

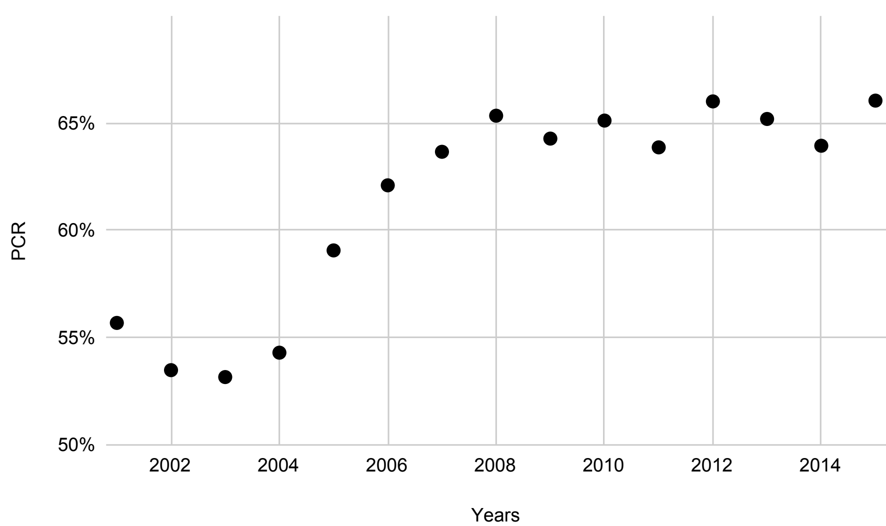

Figure 5.8: *The Phd Completion Rates PCR_n applied at the scaling factor of $(x = 1)$ to the National Data of 284 graduate institutions in each year n starting in 2001.*

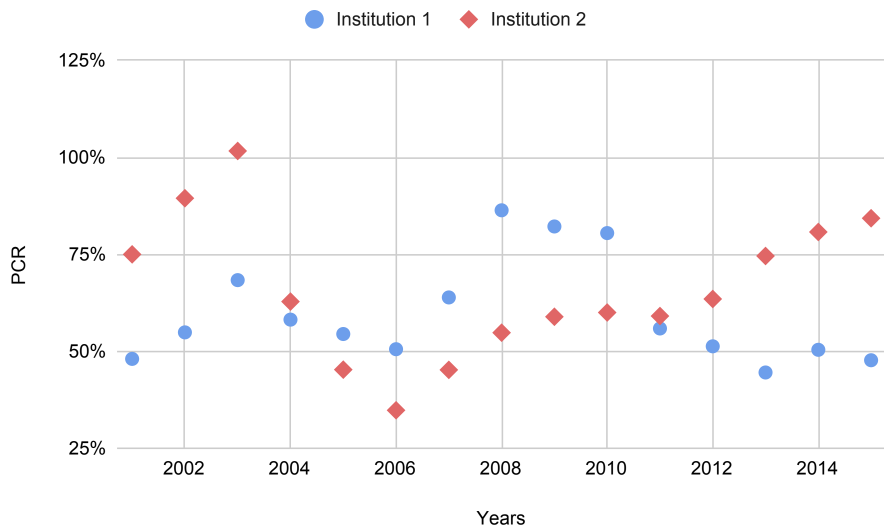

Figure 5.9: *The PhD Completion Rates PCR_n applied at the scaling factor of $(x = 1)$ to two PhD-granting institutions will complete data spaces in each year n starting in 2001.*

National and Institutional Levels. In both cases, the use of this measure continues to be intended to find trends in completion rates and not single values, and additionally, does not appear to follow the trend of the measure at the National Level.

Master's completion rate

The Master's Completion Rate is variable in its application. First, we consider the completion rate of Master's students at the National Level. As the variable F_n does not distinguish between first year students enrolled as Master's or PhD Students, we will only consider the Master's-granting institutions for this part. This is done to limit the effect that the first year enrollments of a PhD-granting institution would have on the results of this measure. The National Level will then be the total data from 70 Master's-granting institutions. Figure 5.10 shows the application of this measure to the National Level data between 2001 and 2019 at a scaling factor of $(x = 1)$. With this measure applied to an institution where a Master's degree is the only graduate degree awarded, we note an overall increase in the percentage of degrees awarded, similar to Figure 5.8.

Considering then the application of this measure to the Institutional Level, we consider two institutions with complete data points among the 70 Master's-granting institutions. Figure 5.11

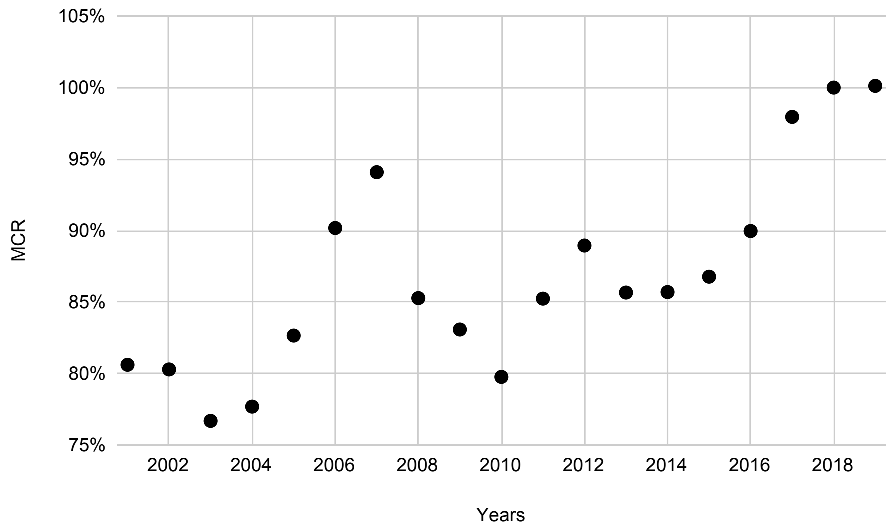

Figure 5.10: *The Master's Completion Rate MCR_n applied at the scaling factor of $(x = 1)$ to the National Data from 70 Master's-granting institutions in each year n starting in 2001*

shows the application of this measure to these two institutions at a scale factor of ($x = 1$), and we find that the application of this measure is different for each institution.

The Master's Completion Rate can be applied to both PhD-granting and Master's-granting institutions. Figure 5.12 shows the application to one of each type of institution with complete data points at a scaling factor of ($x = 1$). As seen above, we find different results for each institution, as expected. Also, the completion rate is notably higher for a Master's-granting institution for the majority of Figure 5.12, which we would expect from an institution where this is the only graduate degree offered. For the PhD-granting institution, the results of the measure are lower, also expected for an institution with multiple graduate degree options. Notably, the two institutions of this figure are not always as separated in completion rates. Again, we would like to highlight the value of determining trends in the data rather than particular points.

5.7 Discussion

The results of this study consider viable measures of retention and graduate completion rates. The measures constructed above consider retention and completion from the perspective of enrollment

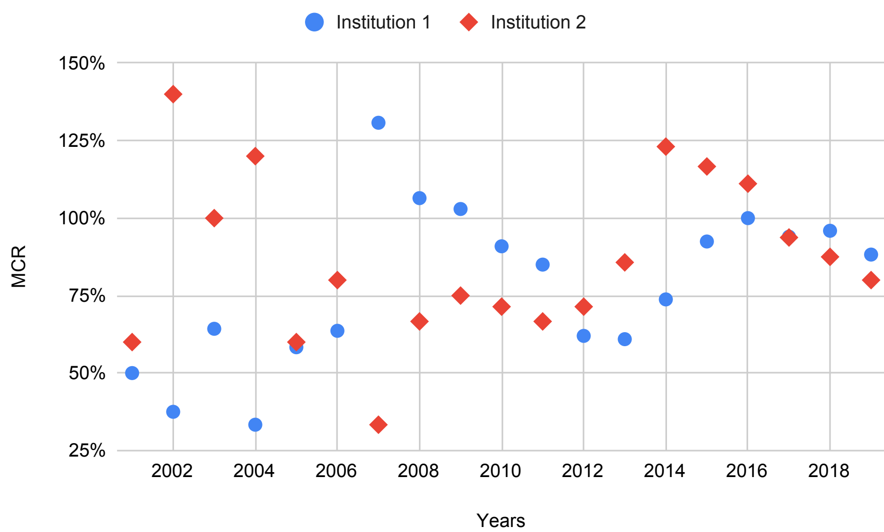

Figure 5.11: *The Master's Completion Rates MCR_n applied at the scaling factor of $(x = 1)$ to two Master's-granting institutions will complete data spaces in each year n starting in 2001.*

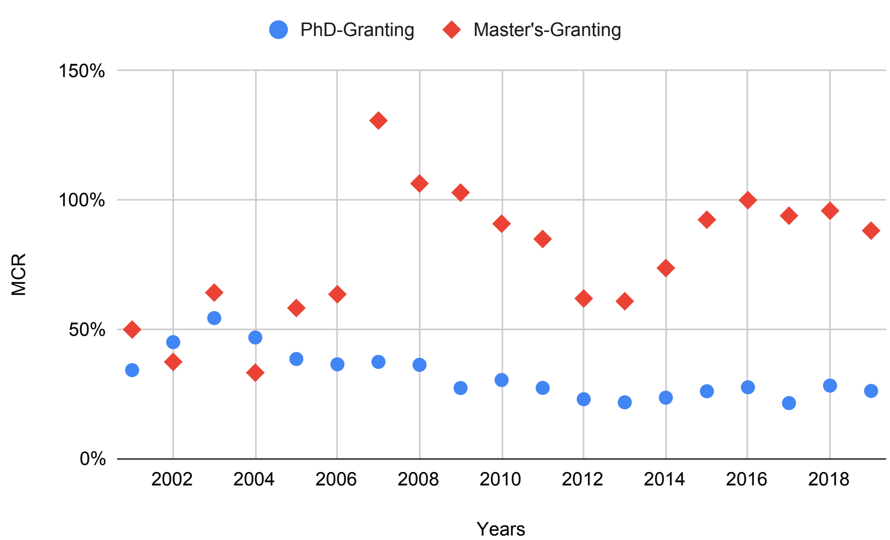

Figure 5.12: *The Master's Completion Rate MCR_n of two graduate institutions, one a PhD-granting and the other a Master's-granting. The scale factor of $(x = 1)$ was applied to both sets between 2001 and 2019.*

data, rather than tracking individual students. In this section, we will situate the results of this study within the context of these previous studies and the insights these individual measures provide for physics graduate programs in the 20 years of survey data from AIP.

5.7.1 Considering trends of measures

The established means of determining retention, persistence, and attrition have traditionally resulted in single values. For example, the seminar works of Lovitts established a retention of graduate students of nearly 50%[40]. Many of these prior studies considered the retention of particular cohorts of graduate students. With many studies presenting values for retention and attrition with no model enumerated, the authors presume that the studies listed in Section **??** considered the cohorts' particular values as directly measured from the students' data. By utilizing survey data from numerous institutions, we developed the measures of this study that could be used to find single values for retention or completion rates for one institution for a particular year.

What we propose instead is that the results of this study be read as a whole, and not as individual values. For example, Figure 5.3 shows the national rate of retention between 2001 and 2022. If we consider one point in this Figure, say GRR_{2020}, the yearly retention rate is 101%. A retention rate over 100% makes little sense, as a way for this to happen is for students to be counted in one year where they were not counted the year before. If we consider the two years surrounding this point, $GRR_{2019} = 95\%$ and $GRR_{2021} = 96\%$, we see the retention rate increase and then drop again. This potentially highlights a significant change in the student population in these years. At this point, it is difficult to determine this level of change in the national numbers of students, where the total population is over 15,000 students. Significant further analysis is required to determine the reasons for this variation.

5.7.2 Troubles with a yearly retention rate

There are some immediate concerns with the results of the Yearly Retention Rate. First that there are numerous points over 100%, particularly in the institutional example of Figure 5.5. The two major variables that dominate the measure are the Total Student Populations for years n and $n-$

1, particularly in the case of the National data, where the Total Student Population is between 9000 and 16000, and the number of PhDs awarded in that same time frame is between 1000 and 2000. With this in mind, we presume that the primary way for this measure to return values over 100% is when the Total Student Population of year n is higher than it is in year $n - 1$, without the corresponding first year enrollment. Considering this phenomenon at the institutional level, the scale of difference between variables is consistent with the National scale. Considering the Institutional Retention rates of Figure 5.5, we find values over 100% as we did with the National Retention Rates. We additionally found each institutional Retention rate to have much more variety than the National Retention Rate.

Considering the National Retention rate as a whole, it is difficult to determine any significant trends in this set of results. This is additionally true for the institutional sets of results. Significantly more quantitative analysis is needed to make sense of the yearly retention rate. As mentioned before, the results of this study should be considered in the context of multiple years. The retention rate measures are constructed using two years of Total Student Populations, and if one of these values is incorrectly reported, then that one incorrect value impacts two rates. This has a particular impact at the institutional level and is likely part of the reason for the numerous jumps in rates that are often seen in the results of this measure. More investigations are needed to make sense of the trends in the rates determined by this measure.

Difficulties with data

Different ways to account for over 100% retention include students transitioning to full-time from part-time, which would only account for the total. Students transferring in should be counted as first years as defined in the AIP solicitation. The solicitation specifies the time frame that each variable needs to be counted for, but it may not be the way each institution interpreted the time frame. As retention rates over 100% are consistent between the Institutional and National Levels, we find it unlikely that misinterpreting the solicitation had a significant effect on the data collected. A particular limitation of this study concerns the survey data from AIP. All of the data for this study is self-reported by each physics graduate program. 22% of the total possible data for this study is unfilled. As mentioned above, a Master's-granting institution would not report PhDs awarded, but only 7 Master's-granting institutions have complete data between 2000 and 2022, and only 67 of the total 284 institutions have full data in this time frame. Further analysis is needed to understand how these measures could be used for institutions that have unfilled data points.

5.7.3 Trends in graduate completion rates

As mentioned before with the Graduate Retention Rate, we wish to present the data on the Graduate Retention Rates as a whole rather than individual rates for particular years. If we consider the scale factor of $(x = 1)$ for each of the Completion Rate figures above, we additionally identify a notable change in some. For example, in Figure 5.8, there is a notable increase in the number of National PhD Completion Rates. In other Figures, it is difficult to determine any significant trends by inspection, such as in Figure 5.9. In all cases of this analysis, considering what trends can be determined from the Completion Rates of this dataset will be avenues for future research.

In their current states, the Completion Rates of this study accomplish a few critical elements. First, as a cursory confirmation of previous literature which includes the consideration of students earning a PhD at a rate of 50-60% [39–45;114]. If we consider Figures 5.8 and 5.9, both tend to show this trend by inspection, particularly in the $x = 1$ case of the National PhD Completion Rate.

Considering then the comparison of Master's Completion Rates between Phd-granting and Master's-granting institutions. The Master's Completion Rate for a PhD-granting institution shows a trend that appears consistent with the models developed in prior literature. If we consider the Master's-granting institution of Figure 5.12, we see a trend that is similar in shape and theme with the PhD Completion Rate of a PhD-Granting Institution. This provides an interesting path for future analysis. If the Master's Completion Rate of a Master's-granting institution (where the Master's is the only degree awarded) is not on the same scale of the PhD Completion Rate of a PhD-granting institution (where the PhD is not the only degree awarded) then what relationship do these two completion rates have with one another. Furthermore, what can both of these rates say about the overall completion rate of graduate students in physics?

If we consider the measures in context with one another, we developed two measures of graduate degree completion. The PhD and Master's Completion Rates are both comparisons of degrees awarded with the first year populations, but utilizing both measures to get a sense of the graduate completion rates for physics students is difficult. Some immediate difficulties in considering the two measures together are the alignment of time scales. Both measures consider a time scale in terms of when graduate degrees are expected to be awarded, but more work will be required to determine the best way to align these completion rates. This will be especially true for considering these measures on the national scale.

5.7.4 Resolving measures for institutional data

These measures were designed to measure graduate student retention and completion rates, with no particular emphasis on population sizes. We made no adjustments to the measures in applying them to the institutional data. We presume that the national versions of each measure could show a more complete picture of the retention and completion rates of Physics Graduate Education. In particular, for the Completion Rates, the scale factor $x = 1$ includes more opportunities for graduate students to be awarded degrees, so we consider each use of this scale factor to be a potentially more complete picture of the rate of degrees awarded. In transitioning the measures to the institutional level, this means that the use of this scale factor may also provide a clearer picture of an institution's completion rate between 2001 and 2017.

Transitioning this work to the institutional level provides several opportunities for further analysis. First, to determine what the Retention Rate of an institution is showing. The issue of over 100% is consistent across all institutions and seen at the National level. More analysis will be necessary to make sense of the changes in student population that are more varied at the institutional level than is muted at the National level. A second avenue of further analysis is the relationship between the institutional PhD and Master's completion rates.

Chapter 6

Conclusions and future work

6.1 Group epistemic frames of value-laden discussions

We identified 3 group epistemic frames that faculty engage in during value-laden discussions: Facilitator, Sharing Science, and Arguing Claims Frames. The process of identifying these group frames was adapted from well-established methodologies of frames and framing literature. To apply this framework to a context of faculty value-laden discussions, we consider elements of argumentation theory and values to further investigate the ways that faculty collectively make sense of the discussions of the fellowship. The 3 group frames are similar to established frames in the literature. The Sharing Science and Arguing Claims epistemic frames have higher engagement levels from the participants. The warrants and values invoked by the participants in these two frames suggest two ways that the faculty consider the value-laden discussions, and help to inform avenues for further investigations into how the participants shift into these group frames.

6.2 The structures of physics graduate education

We have identified 10 structures of physics graduate education, based on public data gathered from 181 PhD-granting institutions: Core Courses, Elective Courses, Comprehensive Content Exams, Research Readiness Exams, the Dissertation, Supervisory Committee, Annual Reports, Training, the TA Requirement, and Research Presentations. We have identified the structures that are more common in physics graduate education as well as those less common. The Core and Elective Courses and the two Candidacy Exam structures are commonly required by programs, yet there is substantial variety in the way these programs are operationalized. The Dissertation is the only universally required structure, and universally vague about providing information about the programmatic expectations, including the writing, preparation, and defense. It is also a structure that has not changed significantly in the way it is described since it was implemented in 1861. The Supervisory Committee spans two major milestones: candidacy exams and the dissertation. The structure itself focuses on the faculty demographics of the group, and less on its function and purpose. Among the less common structures in physics graduate education, we did find many programs that specify requirements for the types and formats of training required for graduate students. Requiring graduate students to be teaching assistants is not a common structure, with only 25% of programs requiring students to be a TA at least once. The annual report is an uncommon structure, and vague about the way the student is expected to report their progress each year to their committee. It is additionally one of the few structures present between the student's candidacy and dissertation defense. Research presentation is the least common structure, with very few institutions requiring students to present their research in places potentially outside their research group, potentially giving those students valuable experience in discussing their work.

The structures of this study additionally present a narrative of physics graduate education that considers multiple structures. First, graduate programs strongly emphasize the learning of physics content knowledge in the first few years of graduate education before students transition into a researcher role until their defense. Physics graduate education additionally prepares students in physics content knowledge and assesses them in it (sometimes multiple times), but only assesses them in research capacity and knowledge, with very few opportunities to prepare them in gaining it. The present goals of candidacy exams seem to be split, and this divide may be influencing negative impressions of the candidacy process for both students and faculty.

We propose several recommendations for physics graduate education based on this work:

- Physics graduate programs should articulate how requirements align with their learning goals to both students and faculty.

- Departments should ensure that structures are in place to prepare students for significant milestones.
- Departments should evaluate their doctoral curriculum to identify redundancies where structures may not be supporting one another.

6.3 Graduate student retention and completion rates

We developed several measures of retention and completion rates for graduate students. The survey data from AIP provided years of enrollment and degrees awarded data for the graduate programs in the United States. We first introduced the graduate yearly retention rate GRR_n. We additionally propose an alternative to this measure that only considers the retention of PhD students, as PRR_n. The second half of this study created measures for calculating the completion rates of first-year physics graduate students for both PhDs and Master's Degrees as a Degree Completion Rate (DCR_n), a PhD Completion Rate (PCR_n), and a Master's Completion Rate (MCR_n). The intention was to build these measures to be applied to the national populations of each year (i.e., the national level) and each institution (i.e., the institutional level).

We propose that the value in these results is found in considering the rates across years, not as individual points. Of particular interest are the moments when the retention rate is over 100%. These moments are present on both the National and Institutional Levels, and are not particular to specific years in this dataset. Making sense of overall trends in this data and ensuring that the measure is correct are avenues of further exploration. Considering the graduate completion rates measures, we identify trends in each measure at the National Scale, consistent with trends previously identified. At the institutional level, the measures return values that are on the same scale as the national level, but not in appearance. This suggests that each completion rate at the institutional level is dependent on the context and flux in populations for that particular institution, and also should be considered as a trend and not individual rates.

6.4 Future Work

6.4.1 Group epistemic frames of value-laden discussions

The study presented in this dissertation highlights an opportunity to investigate the presence of epistemic frames of faculty discourse in different contexts. In this study, we focused on the in-person discussions of the fellowship, and the transition into a virtual environment provides a host of possibilities for further work, including whether the criteria for identifying epistemic frames change during this transition, to whether the frames identified above would maintain the same nature in a virtual environment.

The frames determined through this study emerged from the types of engagement from the participants, and in this way, they are intrinsic to this particular fellowship. It will be particularly interesting then to see the interaction of these frames within an entirely different context, say the problem-solving environment studied in previous frames literature. In those contexts, there is a clear task for the students to engage in, and the epistemic frames identified are closely tied to the general task being performed. One might argue that the fellowship discussions are similarly task-oriented and follow the same established structure from the previous work of epistemic frames in PER. At this time, we would require more information about group discussion dynamics to fully answer this question.

Another avenue this study has is further investigating the intersection of epistemic frames and more ethically minded discussions. Especially in how certain prompts from a moderator influence the epistemic frame that participants engage in, and what helps to transition someone between different e-frames. Ultimately, this path can help understand first the frames that highlight these ethically minded discussions in consistent ways that transition discussions toward these types of frames that promote higher levels of engagement. From the segments observed during the Epistemic Frame B, we see the participants engaging with the discourse through more personal anecdotes, and now the work shifts to seeing what impact that frame has on the participants' ethical decision-making. Ultimately, by understanding how faculty scientists respond to varying value-laden discussions, we can potentially start to see how to prompt those scientists to be in particular epistemic frames.

6.4.2 The structures of physics graduate education

This work serves as a beginning step in identifying the structures of a graduate degree in physics. Similar work must also be done for the numerous ways that students can earn a physics Master's degree. The architecture, in particular, of this study should prove optimal to continue this work towards including more degree types, such as Master's, but also dual degree types, such as those receiving a PhD in Physics and Engineering, Computer Science, etc.

Within the study itself, there are a number of different studies to be done concerning this data. In particular, identifying patterns in categories of groups of institutions, including any consistent types of requirements, and whether there are any additional metrics to compare these institutions with. These comparative measures include higher student retention or particular areas of physics sub-fields discussed at those institutions. Doing so may also help with understanding how program requirements are formed or changed. Another avenue could be conducting a path analysis of program structures and how the structures identified may inform each other.

In the framework of a structure-deficit model, the structures identified in this study each impact the way that students progress through physics graduate programs. Further analysis is needed to understand more fully the impact that the variety of approaches in each structure has on students progressing in programs. Doing so can start to address questions of conformity of practices and what approaches best support graduate students. Looking additionally at furthering specific aspects within the graduate program itself, this study helps to highlight numerous paths of further investigation. With the existing extensive literature of candidacy exams, further analysis into the variations of why departments choose particular combinations of exam types or expectations of those combinations is one avenue. Additionally, the writing and preparation of the dissertation from the perspective of faculty and students could provide some valuable insight as to why this requirement is so limited in this study.

6.5 Graduate student retention and completion rates

There are multiple opportunities for advancing this work. To connect this work with the study of identified program structure in Chapter 4, the natural path forward is to adjust the measures developed here to be applied to the 284 institutions individually. The first measure requires some significant work to better capture how the retention of graduate students is changing each year before it can start to be applied on the institutional level. The completion rate measures are in a better position; however, care must be taken to ensure that the measures can be scaled to the varying institution sizes. Ensuring that the data that was used for this study is better understood for the varying time scales, and that the measures can comfortably cover years of unreported data is critical for furthering this work. Connecting the changing graduate student population and retention measures with each institution's graduate requirements, we can begin to contribute to programs investigating potential institutional practices that contribute to retention and completion rates. This study also provides a similar foundation to Lovitt's work, where a combination of this data set with qualitative analysis can further the understanding of current physics graduate student experiences.

An additional element this group is interested in is broadening the scale of the measures described in this work. The AIP survey is specific to physics degrees awarded and student enrollment. Other organizations collect these types of data from broader disciplines or categorizations of institutions. Further analysis will be needed to see if the measures described in this study could be applied to a broader scale. Another element of further analysis is to connect this effort to standardize the measures of retention and completion rates to path analysis as described above. Doing so can help to highlight areas of further analyses, such as admissions, to determine more consistent ways to track the progress of graduate students.

Bibliography

[1] Synthesis of discipline-based education research in physics | Phys. Rev. ST Phys. Educ. Res., . URL https://journals.aps.org/prper/abstract/10.1103/PhysRevSTPER.10.020119.

[2] *The International Handbook of Physics Education Research: Special Topics*. AIP Publishing LLC, March 2023. ISBN 978-0-7354-2551-4. doi: 10.1063/9780735425514. URL https://pubs.aip.org/books/monograph/161/The-International-Handbook-of-Physics-Education.

[3] *Read "Fostering Integrity in Research" at NAP.edu*. . doi: 10.17226/21896. URL https://www.nap.edu/read/21896/chapter/20.

[4] Alan Leshner and Layne Scherer, editors. *Graduate STEM Education for the 21st Century*. National Academies Press, Washington, D.C., August 2018. ISBN 978-0-309-47273-9. doi: 10.17226/25038. URL https://www.nap.edu/catalog/25038.

[5] *The Path to the Ph.D.: Measuring Graduate Attrition in the Sciences and Humanities*. National Academies Press, Washington, D.C., January 1996. ISBN 978-0-309-05482-9. doi: 10.17226/5195. URL http://www.nap.edu/catalog/5195.

[6] *Read "Doctorate Recipients From United States Universities, 1958-1966: Sciences, Humanities, Professions, Arts: A Statistical Report" at NAP.edu*. . doi: 10.17226/20712. URL https://nap.nationalacademies.org/read/20712/chapter/1.

[7] Kate Kirby and Frances A. Houle. Ethics and the Welfare of the Physics Profession. *Physics Today*, 57(11):42–46, November 2004. ISSN 0031-9228. doi: 10.1063/1.1839376. URL https://physicstoday.scitation.org/doi/10.1063/1.1839376. Publisher: American Institute of Physics.

[8] Yu Xie, Kai Wang, and Yan Kong. Prevalence of Research Misconduct and Questionable Research Practices: A Systematic Review and Meta-Analysis. *Science and Engineering Ethics*, 27(4):41, June 2021. ISSN 1471-5546. doi: 10.1007/s11948-021-00314-9.

[9] Nicholas H. Steneck and Ruth Ellen Bulger. The history, purpose, and future of instruction in the responsible conduct of research. *Academic Medicine: Journal of the Association of American Medical Colleges*, 82(9):829–834, September 2007. ISSN 1040-2446. doi: 10.1097/ACM.0b013e31812f7d4d.

[10] National Academy of Sciences, National Academy of Engineering, and and Institute of Medicine. *On Being a Scientist: A Guide to Responsible Conduct in Research: Third Edition*. March 2009. ISBN 978-0-309-11970-2. doi: 10.17226/12192. URL https://nap.nationalacademies.org/catalog/12192/on-being-a-scientist-a-guide-to-responsible-conduct-in.

[11] 19.1 Guidelines on Ethics (Full Statement), . URL http://www.aps.org/policy/statements/guidlinesethics.cfm.

[12] Trisha Phillips, Franchesca Nestor, Gillian Beach, and Elizabeth Heitman. America COMPETES at 5 years: An Analysis of Research-Intensive Universities' RCR Training Plans. *Science and Engineering Ethics*, 24(1):227–249, February 2018. ISSN 1471-5546. doi: 10.1007/s11948-017-9883-5. URL https://doi.org/10.1007/s11948-017-9883-5.

[13] Frances A. Houle, Kate P. Kirby, and Michael P. Marder. Ethics in physics: The need for culture change. *Physics Today*, 76(1):28–35, January 2023. ISSN 0031-9228. doi: 10.1063/PT.3.5156. URL https://doi.org/10.1063/PT.3.5156.

[14] James M. DuBois, Jeffrey M. Dueker, Emily E. Anderson, and Jean Campbell. The development and assessment of an NIH-funded research ethics training program. *Academic Medicine: Journal of the Association of American Medical Colleges*, 83(6):596–603, June 2008. ISSN 1938-808X. doi: 10.1097/ACM.0b013e3181723095.

[15] Michael D. Mumford, Lynn D. Devenport, Ryan P. Brown, Shane Connelly, Stephen T. Murphy, Jason H. Hill, and Alison L. Antes. ARTICLES: Validation of Ethical Decision Making Measures: Evidence for a New Set of Measures. *Ethics & Behavior*, 16(4):319–345, October 2006. ISSN 1050-8422. doi: 10.1207/s15327019eb1604_4. URL https://doi.org/10.1207/s15327019eb1604_4. Publisher: Routledge _eprint: https://doi.org/10.1207/s15327019eb1604_4.

[16] James F. Johnson, Zhanna Bagdasarov, Shane Connelly, Lauren Harkrider, Lynn D. Devenport, Michael D. Mumford, and Chase E. Thiel. Case-based ethics education: the impact of cause complexity and outcome favorability on ethicality. *Journal of empirical research on human research ethics: JERHRE*, 7(3):63–77, July 2012. ISSN 1556-2654. doi: 10.1525/jer.2012.7.3.63.

[17] Daniel M. Hausman. *Preference, Value, Choice, and Welfare*. Cambridge University Press, Cambridge, 2011. ISBN 978-1-107-01543-2. doi: 10.1017/CBO9781139058537. URL https://www.cambridge.org/core/books/preference-value-choice-and-welfare/1406E7726CE93F4F4E06D752BF4584A2.

[18] Heather Douglas. Inductive Risk and Values in Science. *Philosophy of Science*, 67(4):559–579, 2000. ISSN 0031-8248. URL https://www.jstor.org/stable/188707. Publisher: [The University of Chicago Press, Philosophy of Science Association].

[19] Michael J. Reiss. Teaching Ethics in Science. *Studies in Science Education*, 34(1):115–140, January 1999. ISSN 0305-7267. doi: 10.1080/03057269908560151. URL https://doi.org/10.1080/03057269908560151. Publisher: Routledge _eprint: https://doi.org/10.1080/03057269908560151.

[20] Matthew J. Brown. Values in Science beyond Underdetermination and Inductive Risk. *Philosophy of Science*, 80(5):829–839, 2013. ISSN 0031-8248. doi: 10.1086/673720. URL https://www.jstor.org/stable/10.1086/673720. Publisher: [The University of Chicago Press, Philosophy of Science Association].

[21] Tyler Garcia. *Investigating Changes in Scientists' Ethical Decision Making and Course Design*. PhD thesis, ProQuest LLC, 2024. ISBN: 9798382779317 Publication Title: ProQuest LLC ERIC Number: ED657210.

[22] Thomas J. Bing and Edward F. Redish. Analyzing problem solving using math in physics: Epistemological framing via warrants. *Physical Review Special Topics - Physics Education Research*, 5(2):020108, December 2009. doi: 10.1103/PhysRevSTPER.5.020108. URL https://link.aps.org/doi/10.1103/PhysRevSTPER.5.020108. Publisher: American Physical Society.

[23] Deepa Chari, Hai Nguyen, Dean Zollman, and Eleanor Sayre. Student and instructor framing in upper-division physics. *American Journal of Physics*, 87, April 2017. doi: 10.1119/1.5120392.

[24] Rachel E. Scherr and David Hammer. Student Behavior and Epistemological Framing: Examples from Collaborative Active-Learning Activities in Physics. *Cognition and Instruction*, 27(2):147–174, April 2009. ISSN 0737-0008. doi: 10.1080/07370000902797379. URL https://doi.org/10.1080/07370000902797379. Publisher: Routledge _eprint: https://doi.org/10.1080/07370000902797379.

[25] Paul W. Irving, Mathew Sandy Martinuk, and Eleanor C. Sayre. Transitions in students' epistemic framing along two axes. *Physical Review Special Topics - Physics Education Research*, 9(1):010111, April 2013. doi: 10.1103/PhysRevSTPER.9.010111. URL https://link.aps.org/doi/10.1103/PhysRevSTPER.9.010111. Publisher: American Physical Society.

[26] THOMAS LLOYD MALONE. *A History of the Doctor of Philosophy Dissertation in the United States 1861-1930*. Ph.D., Wayne State University, United States – Michigan, 1981. URL https://www.proquest.com/docview/303183103/abstract/D13FFBDCDC594796PQ/1. ISBN: 9798662056459.

[27] Idaykis Rodriguez, Renee Michelle Goertzen, Eric Brewe, and Laird H. Kramer. Developing

a physics expert identity in a biophysics research group. *Physical Review Special Topics - Physics Education Research*, 11(1):010116, June 2015. doi: 10.1103/PhysRevSTPER.11.010116. URL https://link.aps.org/doi/10.1103/PhysRevSTPER.11.010116. Publisher: American Physical Society.

[28] Raymond G. Greene, Barry J. Hardy, and Steven J. Smith. Graduate Education: Adapting to Current Realities. *Issues in Science and Technology*, 12(2):59–66, 1995. ISSN 0748-5492. URL https://www.jstor.org/stable/43311512. Publisher: University of Texas at Dallas.

[29] John Guillory. The System of Graduate Education. *PMLA*, 115 (5):1154–1163, October 2000. ISSN 0030-8129, 1938-1530. doi: 10.2307/463294. URL https://www.cambridge.org/core/journals/pmla/article/abs/system-of-graduate-education/6647B460FACD47B3D88234C5A9A610E9.

[30] Chris M. Golde and Timothy M. Dore. At Cross Purposes: What the Experiences of Today's Doctoral Students Reveal about Doctoral Education. Technical report, 2001. URL https://eric.ed.gov/?id=ED450628. ERIC Number: ED450628.

[31] P. V. August, J.m. Swift, D. Q. Kellogg, G. Page, P. Nelson, J. Opaluch, J. S. Cobb, C. Foster, and A.j. Gold. The T Assessment Tool: A Simple Metric for Assessing Multidisciplinary Graduate Education. *Journal of Natural Resources and Life Sciences Education*, 39(1):15–21, 2010. ISSN 1539-1582. doi: 10.4195/jnrlse.2009.0004. URL https://onlinelibrary.wiley.com/doi/abs/10.4195/jnrlse.2009.0004. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.4195/jnrlse.2009.0004.

[32] Heidi Estrem and Brad E. Lucas. Embedded Traditions, Uneven Reform: The Place of the Comprehensive Exam in Composition and Rhetoric PhD Programs. *Rhetoric Review*, 22(4):396–416, October 2003. ISSN 0735-0198. doi: 10.1207/S15327981RR2204_4. URL https://doi.org/10.1207/S15327981RR2204_4. Publisher: Routledge _eprint: https://doi.org/10.1207/S15327981RR2204_4.

[33] Joel Fajans, J. Garcia, Jim Gates, Allen Goldman, Peter Jung, and Michael Paesler. Report of the Joint AAPT-APS Task Force.

[34] Sean M. Bell, Jessica Blumstein, Katja Brose, Adam Carroll, Jean Chang, Julia Charles, Elizabeth S. Haswell, Melissa Michelitsch, Julia C. Owens, Christopher K. Patil, Rebecca Smith, Jon Tupy, Emily Walsh, and Tracy Ware. Defining success in graduate school. *Molecular Biology of the Cell*, 25(13):1942–1944, July 2014. ISSN 1059-1524. doi: 10.1091/mbc.e14-03-0793. URL https://www.molbiolcell.org/doi/full/10.1091/mbc.e14-03-0793. Publisher: American Society for Cell Biology (mboc).

[35] Mary K. Enright and Drew Gitomer. Toward a Description of Successful Graduate Students. *ETS Research Report Series*, 1989(1):i–37, 1989. ISSN 2330-8516. doi: 10.1002/j.2330-8516.1989.tb00335.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/j.2330-8516.1989.tb00335.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.2330-8516.1989.tb00335.x.

[36] Rodney T. Hartnett and Warren W. Willingham. The Criterion Problem: What Measure of Success in Graduate Education? *Applied Psychological Measurement*, 4(3):281–291, July 1980. ISSN 0146-6216. doi: 10.1177/014662168000400301. URL https://doi.org/10.1177/014662168000400301. Publisher: SAGE Publications Inc.

[37] Clair Hughes and Simon Barrie. Influences on the assessment of graduate attributes in higher education. *Assessment & Evaluation in Higher Education*, 35 (3):325–334, May 2010. ISSN 0260-2938. doi: 10.1080/02602930903221485. URL https://doi.org/10.1080/02602930903221485. Publisher: Routledge _eprint: https://doi.org/10.1080/02602930903221485.

[38] Warren W. Willingham. Predicting Success in Graduate Education. *Science*, 183(4122): 273–278, January 1974. doi: 10.1126/science.183.4122.273. URL https://www.science.org/doi/abs/10.1126/science.183.4122.273. Publisher: American Association for the Advancement of Science.

[39] Chris M. Golde. The Role of the Department and Discipline in Doctoral Student Attrition: Lessons from Four Departments. *The Journal of Higher Education*, 76(6):669–700, November 2005. ISSN 0022-1546. doi: 10.1080/00221546.2005.11772304. URL https://doi.org/10.1080/00221546.2005.11772304. Publisher: Routledge _eprint: https://doi.org/10.1080/00221546.2005.11772304.

[40] Barbara E. Lovitts. *Leaving the Ivory Tower: The Causes and Consequences of Departure from Doctoral Study*. Rowman & Littlefield Publishers, Inc, 2001. ISBN 978-0-7425-0942-9. ERIC Number: ED462026.

[41] Steven Terrell, Martha Snyder, Laurie Dringus, and Elizabeth Maddrey. A Grounded Theory of Connectivity and Persistence in a Limited Residency Doctoral Program. *The Qualitative Report*, 17(31):1–14, July 2012. ISSN 1052-0147. doi: 10.46743/2160-3715/2012.1747. URL https://nsuworks.nova.edu/tqr/vol17/iss31/2.

[42] Cathy Wendler, Brent Bridgeman, Fred Cline, Catherine Millett, JoAnn Rock, Nathan Bell, and Patricia McAllister. The Path Forward: The Future of Graduate Education in the United States. Technical report, Educational Testing Service, April 2010. URL https://eric.ed.gov/?id=ED509441. Publication Title: Educational Testing Service ERIC Number: ED509441.

[43] Ph D. Michael T. Nettles and Ph D. Catherine M. Millett. *Three Magic Letters*. Johns Hopkins University Press, 2006. ISBN 978-0-8018-8232-6. doi: 10.56021/9780801882326. URL https://www.press.jhu.edu/books/title/3353/three-magic-letters.

[44] Roger Geiger. Research, graduate education, and the ecology of American universities: an interpretive history. In Bjorn Wittrock and Sheldon Rothblatt, editors, *The European and American University since 1800*, pages 234–260. Cambridge University Press, Cambridge, 1993. ISBN 978-0-521-43165-1. doi: 10.1017/CBO9780511720925.007. URL https://www.cambridge.org/core/books/european-and-american-university-since-1800/

research-graduate-education-and-the-ecology-of-american-universities-a
68CCD43F17EDC02D509F411680A70A53.

[45] Dropouts from College., . URL https://awspntest.apa.org/record/2006-10217-019.

[46] Nicholas T. Young, N. Verboncoeur, Dao Chi Lam, and Marcos D. Caballero. Rubric-based holistic review represents a change from traditional graduate admissions approaches in physics. *Physical Review Physics Education Research*, 19(1):010134, May 2023. doi: 10.1103/PhysRevPhysEducRes.19.010134. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.19.010134. Publisher: American Physical Society.

[47] Shiva Basir and Eric Burkholder. Investigating faculty perspectives on written qualifying exams in physics. *Physical Review Physics Education Research*, 20(1):010139, May 2024. ISSN 2469-9896. doi: 10.1103/PhysRevPhysEducRes.20.010139. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.20.010139.

[48] Mike Verostek, Casey W. Miller, and Benjamin M. Zwickl. Modeling when and how physics Ph.D. students search for a research group: The role of interests and prior research experiences in timely group integration. *Physical Review Physics Education Research*, 20(2):020119, September 2024. doi: 10.1103/PhysRevPhysEducRes.20.020119. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.20.020119. Publisher: American Physical Society.

[49] John P. Bean. Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education*, 12(2):155–187, June 1980. ISSN 1573-188X. doi: 10.1007/BF00976194. URL https://doi.org/10.1007/BF00976194.

[50] Joseph B. Berger and Jeffrey F. Milem. The Role of Student Involvement and Perceptions of Integration in a Causal Model of Student Persistence. *Research in Higher Education*, 40(6):641–664, 1999. ISSN 0361-0365. URL https://www.jstor.org/stable/40196897. Publisher: Springer.

[51] Cullen F. Goenner and Sean M. Snaith. Predicting Graduation Rates: An Analysis of Student and Institutional Factors at Doctoral Universities. *Journal of College Student Retention*, 5 (4):409–420, February 2004. ISSN 1521-0251. doi: 10.2190/LKJX-CL3H-1AJ5-WVPE. URL https://doi.org/10.2190/LKJX-CL3H-1AJ5-WVPE. Publisher: SAGE Publications Inc.

[52] Lindsay M. Owens, Benjamin M. Zwickl, Scott V. Franklin, and Casey W. Miller. Misaligned Visions for Improving Graduate Diversity: Student Characteristics vs. Systemic/Cultural Factors. December 2018. ISBN 978-1-931024-34-1. URL https://www.per-central.org/items/detail.cfm?ID=14834. ISSN: 2377-2379.

[53] Donald S. Kornfeld. Perspective: research misconduct: the search for a remedy. *Academic Medicine: Journal of the Association of American Medical Colleges*, 87(7):877–882, July 2012. ISSN 1938-808X. doi: 10.1097/ACM.0b013e318257ee6a.

[54] James M. Dubois and Jeffrey M. Dueker. Teaching and Assessing the Responsible Conduct of Research: A Delphi Consensus Panel Report. *The Journal of Research Administration*, 40(1):49–70, 2009. ISSN 1539-1590.

[55] Justin B. Biddle and Rebecca Kukla. *The Geography of Epistemic Risk*. Oxford University Press. ISBN 978-0-19-046775-3. URL https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780190467715.001.0001/acprof-9780190467715-chapter-11. Publication Title: Exploring Inductive Risk Section: Exploring Inductive Risk.

[56] Aidan C Cairns, Caleb Linville, Tyler Garcia, Bill Bridges, Scott Tanona, Jonathan Herington, and James T Laverty. A phenomenographic study of scientists' beliefs about the causes of scientists' research misconduct. *Research Ethics*, page 17470161211042658, September 2021. ISSN 1747-0161. doi: 10.1177/17470161211042658. URL https://doi.org/10.1177/17470161211042658. Publisher: SAGE Publications Ltd.

[57] Paul Braunschweiger and Kenneth W. Goodman. The CITI program: an international online

resource for education in human subjects protection and the responsible conduct of research. *Academic Medicine: Journal of the Association of American Medical Colleges*, 82(9):861–864, September 2007. ISSN 1040-2446. doi: 10.1097/ACM.0b013e31812f7770.

[58] Lauren N. Harkrider, Alexandra E. MacDougall, Zhanna Bagdasarov, James F. Johnson, Chase E. Thiel, Michael D. Mumford, Shane Connelly, and Lynn D. Devenport. Structuring Case-Based Ethics Training: How Comparing Cases and Structured Prompts Influence Training Effectiveness. *Ethics & Behavior*, 23(3):179–198, May 2013. ISSN 1050-8422. doi: 10.1080/10508422.2012.728470. URL https://doi.org/10.1080/10508422.2012.728470. Publisher: Routledge _eprint: https://doi.org/10.1080/10508422.2012.728470.

[59] Whitney Helton-Fauth, Blaine Gaddis, Ginamarie Scott, Michael Mumford, Lynn Devenport, Shane Connelly, and Ryan Brown. A new approach to assessing ethical conduct in scientific work. *Accountability in Research*, 10(4):205–228, December 2003. ISSN 0898-9621. doi: 10.1080/714906104.

[60] Alison L. Antes, Stephen T. Murphy, Ethan P. Waples, Michael D. Mumford, Ryan P. Brown, Shane Connelly, and Lynn D. Devenport. A Meta-Analysis of Ethics Instruction Effectiveness in the Sciences. *Ethics & behavior*, 19(5):379–402, September 2009. ISSN 1050-8422. doi: 10.1080/10508420903035380. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2762211/.

[61] Lauren N. Harkrider, Chase E. Thiel, Zhanna Bagdasarov, Michael D. Mumford, James F. Johnson, Shane Connelly, and Lynn D. Devenport. Improving Case-Based Ethics Training with Codes of Conduct and Forecasting Content. *Ethics & Behavior*, 22(4):258–280, July 2012. ISSN 1050-8422. doi: 10.1080/10508422.2012.661311. URL https://doi.org/10.1080/10508422.2012.661311. Publisher: Routledge _eprint: https://doi.org/10.1080/10508422.2012.661311.

[62] Edward G. Carmines and James A. Woods. Reliability Assessment. In Kimberly Kempf-Leonard, editor, *Encyclopedia of Social Measurement*, pages 361–365. Elsevier,

New York, January 2005. ISBN 978-0-12-369398-3. doi: 10.1016/B0-12-369398-5/00198-5. URL https://www.sciencedirect.com/science/article/pii/B0123693985001985.

[63] Michael Kalichman. Rescuing responsible conduct of research (RCR) education. *Accountability in Research*, 21(1):68–83, 2014. ISSN 1545-5815. doi: 10.1080/08989621.2013.822271.

[64] Alison L. Antes, Ryan P. Brown, Stephen T. Murphy, Ethan P. Waples, Michael D. Mumford, Shane Connelly, and Lynn D. Devenport. Personality and ethical decision-making in research: the role of perceptions of self and others. *Journal of empirical research on human research ethics: JERHRE*, 2(4):15–34, December 2007. ISSN 1556-2646. doi: 10.1525/jer.2007.2.4.15.

[65] Cheryl K. Beeler, Alison L. Antes, Xiaoqian Wang, Jared J. Caughron, Chase E. Thiel, and Michael D. Mumford. Strategies in Forecasting Outcomes in Ethical Decision-making: Identifying and Analyzing the Causes of the Problem. *Ethics & behavior*, 20(2):110–127, March 2010. ISSN 1050-8422. doi: 10.1080/10508421003595935. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2844648/.

[66] Alison L. Antes, Xiaoqian Wang, Michael D. Mumford, Ryan P. Brown, Shane Connelly, and Lynn D. Devenport. Evaluating the effects that existing instruction on responsible conduct of research has on ethical decision making. *Academic Medicine: Journal of the Association of American Medical Colleges*, 85(3):519–526, March 2010. ISSN 1938-808X. doi: 10.1097/ACM.0b013e3181cd1cc5.

[67] Caleb L. Linville, Aidan C. Cairns, Tyler Garcia, Bill Bridges, Jonathan Herington, James T. Laverty, and Scott Tanona. How Do Scientists Perceive the Relationship Between Ethics and Science? A Pilot Study of Scientists' Appeals to Values. *Science and Engineering Ethics*, 29(3):15, April 2023. ISSN 1471-5546. doi: 10.1007/s11948-023-00429-1. URL https://doi.org/10.1007/s11948-023-00429-1.

[68] Kevin C. Elliott. *Values in Science*. Cambridge University Press, 2022.

[69] Liam Kofi Bright. Du Bois' democratic defence of the value free ideal. *Synthese*, 195 (5):2227–2245, May 2018. ISSN 1573-0964. doi: 10.1007/s11229-017-1333-z. URL https://doi.org/10.1007/s11229-017-1333-z.

[70] Thomas J. Bing and Edward F. Redish. Epistemic complexity and the journeyman-expert transition. *Physical Review Special Topics - Physics Education Research*, 8(1):010105, February 2012. doi: 10.1103/PhysRevSTPER.8.010105. URL https://link.aps.org/doi/10.1103/PhysRevSTPER.8.010105. Publisher: American Physical Society.

[71] Andrea A. diSessa. Toward an Epistemology of Physics. *Cognition and Instruction*, 10 (2/3):105–225, 1993. ISSN 0737-0008. URL https://www.jstor.org/stable/3233725. Publisher: Taylor & Francis, Ltd.

[72] David Hammer. Student resources for learning introductory physics. *American Journal of Physics*, 68(S1):S52–S59, July 2000. ISSN 0002-9505. doi: 10.1119/1.19520. URL https://doi.org/10.1119/1.19520.

[73] Michael C. Wittmann. Using resource graphs to represent conceptual change. *Physical Review Special Topics - Physics Education Research*, 2(2):020105, August 2006. doi: 10.1103/PhysRevSTPER.2.020105. URL https://link.aps.org/doi/10.1103/PhysRevSTPER.2.020105. Publisher: American Physical Society.

[74] M. A. van Rees. Discourse analysis and argumentation theory: The case of television talk. *Journal of Pragmatics*, 39(8):1454–1463, August 2007. ISSN 0378-2166. doi: 10.1016/j.pragma.2007.04.005. URL https://www.sciencedirect.com/science/article/pii/S0378216607000616.

[75] Stephen E Toulmin. The Uses of Argument, Updated Edition. page 259.

[76] Sibel Erduran, Shirley Simon, and Jonathan Osborne. TAPping into Argumentation: Developments in the application of Toulmin's Argument Pattern for studying science discourse. *Science Education*, 88:915–933, November 2004. doi: 10.1002/sce.20012.

[77] Ying Chen, Paul W. Irving, and Eleanor C. Sayre. Epistemic game for answer making in learning about hydrostatics. *Physical Review Special Topics - Physics Education Research*, 9(1):010108, March 2013. doi: 10.1103/PhysRevSTPER.9.010108. URL https://link.aps.org/doi/10.1103/PhysRevSTPER.9.010108. Publisher: American Physical Society.

[78] Edited by Deborah Tannen, editor. *Framing in Discourse*. Oxford University Press, Oxford, New York, October 1993. ISBN 978-0-19-507996-8.

[79] Paul Hutchison and David Hammer. Attending to student epistemological framing in a science classroom. *Science Education*, 94(3):506–524, 2010. ISSN 1098-237X. doi: 10.1002/sce.20373. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sce.20373. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sce.20373.

[80] Brian W. Frank and Rachel E. Scherr. Interactional processes for stabilizing conceptual coherences in physics. *Physical Review Special Topics - Physics Education Research*, 8(2):020101, July 2012. doi: 10.1103/PhysRevSTPER.8.020101. URL https://link.aps.org/doi/10.1103/PhysRevSTPER.8.020101. Publisher: American Physical Society.

[81] Kelli Shar, Rosemary S. Russ, and James T. Laverty. Student epistemological framing on paper-based assessments. *Physical Review Physics Education Research*, 16(2):020101, July 2020. doi: 10.1103/PhysRevPhysEducRes.16.020101. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.16.020101. Publisher: American Physical Society.

[82] On the form of a personal epistemology., . URL https://awspntest.apa.org/record/2001-18187-008.

[83] Gregory Bateson. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. University of Chicago Press, Chicago, IL, April 2000. ISBN 978-0-226-03905-3. URL https://press.uchicago.edu/ucp/books/book/chicago/S/bo3620295.html.

[84] Bahar Modir, John D. Thompson, and Eleanor C. Sayre. Students' epistemological framing in quantum mechanics problem solving. *Physical Review Physics Education Research*, 13 (2):020108, August 2017. doi: 10.1103/PhysRevPhysEducRes.13.020108. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.13.020108. Publisher: American Physical Society.

[85] Lesley Lydell. Assessing outcomes in graduate education. *On the Horizon*, 16(2):107–117, January 2008. ISSN 1074-8121. doi: 10.1108/10748120810874504. URL https://doi.org/10.1108/10748120810874504. Publisher: Emerald Group Publishing Limited.

[86] Román Liera, Aireale J. Rodgers, Lauren N. Irwin, and Julie R. Posselt. Rethinking doctoral qualifying exams and candidacy in the physical sciences: Learning toward scientific legitimacy. *Physical Review Physics Education Research*, 19(2):020110, August 2023. doi: 10.1103/PhysRevPhysEducRes.19.020110. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.19.020110. Publisher: American Physical Society.

[87] John Soloski. On Defining the Nature of Graduate Education. *The Journalism Educator*, 49 (2):4–11, June 1994. ISSN 0022-5517. doi: 10.1177/107769589404900202. URL https://doi.org/10.1177/107769589404900202. Publisher: SAGE Publications.

[88] Benedicta Donkor and Jordan Harshman. Learning Goals and Priorities Identified by an Examination of Chemistry Graduate Handbooks. *Journal of Chemical Education*, 100 (10):3774–3783, October 2023. ISSN 0021-9584. doi: 10.1021/acs.jchemed.3c00062. URL https://doi.org/10.1021/acs.jchemed.3c00062. Publisher: American Chemical Society.

[89] Peggy L. Maki and Nancy A. Borkowski. *The Assessment of Doctoral Education: Emerging Criteria and New Models for Improving Outcomes*. Stylus Publishing, LLC, October 2006. ISBN 978-1-57922-178-2. Publication Title: Stylus Publishing, LLC ERIC Number: ED496319.

[90] 2008 Graduate Education Conference Report, . URL https://www.aps.org/publications/reports/2008-grad-edu-conference.

[91] Ramón Barthelemy, MacKenzie Lenz, Alexis Knaub, Jordan Gerton, and Pearl Sandick. Graduate program reform in one department of physics and astronomy: From tragedy to more progressive policies and an evolving culture. *Physical Review Physics Education Research*, 19(1):010102, January 2023. doi: 10.1103/PhysRevPhysEducRes.19.010102. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.19.010102. Publisher: American Physical Society.

[92] Mike Verostek, Casey W. Miller, and Benjamin Zwickl. Analyzing admissions metrics as predictors of graduate GPA and whether graduate GPA mediates Ph.D. completion. *Physical Review Physics Education Research*, 17(2):020115, September 2021. doi: 10.1103/PhysRevPhysEducRes.17.020115. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.17.020115. Publisher: American Physical Society.

[93] MaryBeth Walpole, Nancy W. Burton, Kamau Kanyi, and Altamese Jackenthal. SELECTING SUCCESSFUL GRADUATE STUDENTS: IN-DEPTH INTERVIEWS WITH GRE® USERS. *ETS Research Report Series*, 2002(1), June 2002. ISSN 2330-8516, 2330-8516. doi: 10.1002/j.2333-8504.2002.tb01875.x. URL https://onlinelibrary.wiley.com/doi/10.1002/j.2333-8504.2002.tb01875.x.

[94] Margaret F. King, editor. *Ph.D. completion and attrition: analysis of baseline demographic data from the Ph.D. Completion Project*. Council of Graduate Schools, Washington, D.C, 2008. ISBN 978-1-933042-18-3. OCLC: ocn252862330.

[95] Gary D. Malaney. Differentiation in graduate education. *Research in Higher Education*,

25(1):82–96, March 1986. ISSN 1573-188X. doi: 10.1007/BF00991880. URL https://doi.org/10.1007/BF00991880.

[96] S. L. Pressey, L. C. Pressey, and Elinor J. Barnes. The Final Ordeal. *The Journal of Higher Education*, 3(5):261–264, 1932. ISSN 0022-1546. doi: 10.2307/1974483. URL https://www.jstor.org/stable/1974483. Publisher: Ohio State University Press.

[97] Bahar Memarian, Sasha Gollish, Kimia Moozeh, Darlee Gerrard, Deborah Tihanyi, and Stacy Costa. Comprehensive Exam Variations and Experiences of Doctoral Candidates in Engineering. In *2019 ASEE Annual Conference & Exposition Proceedings*, page 32530, Tampa, Florida, June 2019. ASEE Conferences. doi: 10.18260/1-2--32530. URL http://peer.asee.org/32530.

[98] Doctoral Candidacy and Qualifying Exams: A Guide to Contexts, Costs for Equity, and Possibilities of Transformation, . URL https://pullias.usc.edu/publication/doctoral-candidacy-and-qualifying-exams-a-guide-to-contexts-costs-for-

[99] Joseph C. Hermanowicz. Faculty perceptions of their graduate education. *Higher Education*, 72(3):291–305, September 2016. ISSN 1573-174X. doi: 10.1007/s10734-015-9955-x. URL https://doi.org/10.1007/s10734-015-9955-x.

[100] Timothy T. Baldwin and J. Kevin Ford. Transfer of Training: A Review and Directions for Future Research. *Personnel Psychology*, 41(1):63–105, 1988. ISSN 1744-6570. doi: 10.1111/j.1744-6570.1988.tb00632.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-6570.1988.tb00632.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1744-6570.1988.tb00632.x.

[101] Karen M. Collier and Margaret R. Blanchard. Toward a Holistic Understanding of Factors That Support or Inhibit Graduate Student Success. *Trends in Higher Education*, 2(3):389–408, September 2023. ISSN 2813-4346. doi: 10.3390/higheredu2030023. URL https://www.mdpi.com/2813-4346/2/3/23. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.

[102] Jean M. Curtin, Geneva Blake, and Christine Cassagnau. THE CLIMATE FOR WOMEN GRADUATE STUDENTS IN PHYSICS. *Journal of Women and Minorities in Science and Engineering*, 3(1&2), 1997. ISSN 1072-8325, 1940-431X. doi: 10.1615/JWomenMinorScienEng.v3.i1-2.70. URL https://www.dl.begellhouse.com/journals/00551c876cc2f027,3ff09bf71ca3a958,05d8b4b80348ddf9.html. Publisher: Begel House Inc.

[103] Shaun R. Harper. An anti-deficit achievement framework for research on students of color in STEM. *New Directions for Institutional Research*, 2010(148):63–74, 2010. ISSN 1536-075X. doi: 10.1002/ir.362. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/ir.362. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ir.362.

[104] Joel Alejandro Mejia, Renata A. Revelo, Idalis Villanueva, and Janice Mejia. Critical Theoretical Frameworks in Engineering Education: An Anti-Deficit and Liberative Approach. *Education Sciences*, 8(4):158, December 2018. ISSN 2227-7102. doi: 10.3390/educsci8040158. URL https://www.mdpi.com/2227-7102/8/4/158. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[105] David J. Webb and Cassandra A. Paul. Attributing equity gaps to course structure in introductory physics. *Physical Review Physics Education Research*, 19(2):020126, September 2023. doi: 10.1103/PhysRevPhysEducRes.19.020126. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.19.020126. Publisher: American Physical Society.

[106] Richard R. Valencia, editor. *The Evolution of Deficit Thinking: Educational Thought and Practice*. Routledge, London, November 2012. ISBN 978-0-203-04658-6. doi: 10.4324/9780203046586.

[107] Sehoya Cotner and Cissy J. Ballen. Can mixed assessment methods make biology classes more equitable? *PLOS ONE*, 12(12):e0189610, December 2017. ISSN 1932-6203. doi: 10.1371/journal.pone.0189610. URL https://journals.plos.org/plosone/

article?id=10.1371/journal.pone.0189610. Publisher: Public Library of Science.

[108] Cassandra A. Paul and David J. Webb. Examining equity and graduation rates at two institutions using a course deficit model and the collaborative learning through active sense-making in physics curriculum. *Physical Review Physics Education Research*, 21(1):010107, January 2025. doi: 10.1103/PhysRevPhysEducRes.21.010107. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.21.010107. Publisher: American Physical Society.

[109] *Leaving College*. . URL https://press.uchicago.edu/ucp/books/book/chicago/L/bo3630345.html.

[110] John P. Bean and Barbara S. Metzner. A Conceptual Model of Nontraditional Undergraduate Student Attrition. *Review of Educational Research*, 55(4):485–540, 1985. ISSN 0034-6543. doi: 10.2307/1170245. URL https://www.jstor.org/stable/1170245. Publisher: [Sage Publications, Inc., American Educational Research Association].

[111] Diana Sachmpazidi and Charles Henderson. Departmental support structures for physics graduate students: Development and psychometric evaluation of a self-report instrument. *Physical Review Physics Education Research*, 17(1):010123, March 2021. doi: 10.1103/PhysRevPhysEducRes.17.010123. URL https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.17.010123. Publisher: American Physical Society.

[112] Susan K. Gardner. Student and faculty attributions of attrition in high and low-completing doctoral programs in the United States. *Higher Education*, 58(1):97–112, July 2009. ISSN 1573-174X. doi: 10.1007/s10734-008-9184-7. URL https://doi.org/10.1007/s10734-008-9184-7.

[113] Corinna A. Ethington and John C. Smart. Persistence to graduate education. *Research in Higher Education*, 24(3):287–303, September 1986. ISSN 1573-188X. doi: 10.1007/BF00992076. URL https://doi.org/10.1007/BF00992076.

[114] N. A. Gross, R. Lopez, W. J. Hughes, and Hakeem M. Oluseyi. Theory Into Practice: Applying Lessons Learned From Retention Studies To Build A Diverse Graduate Community. In *AIP Conference Proceedings*, volume 991, pages 110–118, Boston (Massachusetts), 2008. AIP. doi: 10.1063/1.2905128. URL https://pubs.aip.org/aip/acp/article/991/1/110-118/564571. ISSN: 0094243X.

[115] Elaine Seymour and Anne-Barrie Hunter, editors. *Talking about Leaving Revisited: Persistence, Relocation, and Loss in Undergraduate STEM Education*. Springer International Publishing, Cham, 2019. ISBN 978-3-030-25303-5 978-3-030-25304-2. doi: 10.1007/978-3-030-25304-2. URL https://link.springer.com/10.1007/978-3-030-25304-2.

[116] Giorgia Guglielmi. Sexual-minority students more likely to abandon science majors. *Nature*, March 2018. doi: 10.1038/d41586-018-03178-5. URL https://www.nature.com/articles/d41586-018-03178-5. Bandiera_abtest: a Cg_type: News Publisher: Nature Publishing Group Subject_term: Society, Education.

[117] John Darrell Hansen. *Methods for Analyzing Physics Student Retention and Physics Curricula*. PhD, West Virginia University Libraries, January 2023. URL https://researchrepository.wvu.edu/etd/12201.

[118] Bridget M. Maher, Helen Hynes, Catherine Sweeney, Ali S. Khashan, Margaret O'Rourke, Kieran Doran, Anne Harris, and Siun O' Flynn. Medical School Attrition-Beyond the Statistics A Ten Year Retrospective Study. *BMC Medical Education*, 13(1):13, January 2013. ISSN 1472-6920. doi: 10.1186/1472-6920-13-13. URL https://doi.org/10.1186/1472-6920-13-13.

[119] Naomi Karen Kayesa and Maylene Shung-King. The role of document analysis in health policy analysis studies in low and middle-income countries: Lessons for HPA researchers from a qualitative systematic review. *Health Policy OPEN*, 2:100024, December 2021. ISSN 2590-2296. doi: 10.1016/j.hpopen.2020.100024. URL https://www.sciencedirect.com/science/article/pii/S2590229620300228.

[120] Penny Mackieson, Aron Shlonsky, and Marie Connolly. Increasing rigor and reducing bias in qualitative research: A document analysis of parliamentary debates using applied thematic analysis. *Qualitative Social Work*, 18(6):965–980, November 2019. ISSN 1473-3250. doi: 10.1177/1473325018786996. URL https://doi.org/10.1177/1473325018786996. Publisher: SAGE Publications.

[121] Joseph L. Fleiss. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382, 1971. ISSN 1939-1455. doi: 10.1037/h0031619. Place: US Publisher: American Psychological Association.

[122] J. Richard Landis and Gary G. Koch. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1):159–174, 1977. ISSN 0006-341X. doi: 10.2307/2529310. URL https://www.jstor.org/stable/2529310. Publisher: International Biometric Society.

[123] Roster of Physics Departments with Enrollment and Degree Data, 2022 - AIP.ORG, . URL https://ww2.aip.org/statistics/roster-of-physics-departments-with-enrollment-and-degree-data-2022.

[124] Bill Bridges, Caitlin Solis, Joshua Barron, James T. Laverty, Jacquelyn J. Chini, and Rachel Henderson. Investigating the Assessment Landscape of Physics Graduate Programs. pages 64–69, October 2023. ISBN 978-1-931024-39-6. URL https://www.per-central.org/items/detail.cfm?ID=16561. ISSN: 2377-2379.